

**AN EMBODIED COGNITION APPROACH TO
THE ANALYSIS AND DESIGN OF
GENERATIVE AND INTERACTIVE ANIMATION**

A Dissertation
Presented to
The Academic Faculty

by

Ka Nin Chow

In Partial Fulfillment
of the Requirements for the Degree
Ph.D. Digital Media in the
School of Literature, Communication, and Culture

Georgia Institute of Technology
August 2010

**AN EMBODIED COGNITION APPROACH TO
THE ANALYSIS AND DESIGN OF
GENERATIVE AND INTERACTIVE ANIMATION**

Approved by:

Dr. D. Fox Harrell, Advisor
School of Literature, Communication, and
Culture
Georgia Institute of Technology

Dr. Nancy J. Nersessian
School of Interactive Computing
School of Public Policy
Georgia Institute of Technology

Dr. Kenneth J. Knoespel
School of Literature, Communication, and
Culture
Georgia Institute of Technology

Dr. Mark Turner
Case Western Reserve University

Dr. Qi Wang
School of Literature, Communication, and
Culture
Georgia Institute of Technology

Date Approved: May 7, 2010

TABLE OF CONTENTS

	Page
TABLE OF CONTENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF SYMBOLS AND ABBREVIATIONS	xiii
SUMMARY	xv
CHAPTER 1	1
INTRODUCTION	1
Animation as a Medium	1
Animation as a Technology of Illusion	2
Animation as an Art Form	3
Animation in the Digital Age	4
Animation as a Phenomenon of Liveliness	5
CHAPTER 2	7
PROBLEM FORMULATION	7
New Animation Phenomena in the Digital Age	7
Theorizing the New Animation Phenomena	8
Existing Approaches to Animation	10
Definitions of Animation	13
Beyond Single Medium	14

CHAPTER 3	22
THEORETICAL FRAMEWORK	22
The Observer: Perception of Animacy	24
The Image: Understanding Images	37
The Body: Embodied Meaning of Action	50
The Performer: Live Performance of Animation	60
CHAPTER 4	69
NEW PRINCIPLES OF INTERACTIVE GENERATIVE ANIMATION	69
The Observer: Holistic Illusion of Life	72
The Image: Material-Based Imagination	85
The Body: Enduring Interaction	97
The Performer: Computer-Mediated Co-Performance	107
Conclusion	118
CHAPTER 5	120
METHOD OF ANALYSES:	120
THE VARIABLES AND THE TAXONOMY	120
The Variables	120
The Taxonomy	124
CHAPTER 6	129
ANALYSES AND RESULTS	129
The Corpus	130
Conclusion	194

CHAPTER 7	199
DESIGN IMPLICATIONS	199
Motion and Animation Design	200
Interface and Interaction Design	208
Conclusion	213
CHAPTER 8	215
THE GENERATIVE VISUAL RENKU PROJECT	215
Artistic and Theoretical Framework	216
An Example of GVR	220
Summary	231
CHAPTER 9	234
CONCLUSION	234
Animation as an Illusion of Life in the Age of Interactive Computational Generation	235
Animation as a Simulacrum of Life	236
Animation as a Spiritual-Functional Phenomenon	237
REFERENCES	240

LIST OF TABLES

	Page
Table 5.1: The variables and possible values	123
Table 6.1: Variable values taken by the corpus	196

LIST OF FIGURES

	Page
Figure 2.1 An animated “running” green man on a traffic light in Taipei	8
Figure 2.2 Two genies, one in a cartoon animation, another in Macintosh OS X	10
Figure 2.3 Puppeteers being simultaneously performers and audiences of their own show	16
Figure 2.4 A tea-serving karakuri	17
Figure 2.5 Figures on a pacing-horse lamp spinning automatically	18
Figure 2.6 A beholder’s hands and eyes coordinating the panning of a hand-scroll painting (Wu, 1996, p. 58)	19
Figure 3.1 Sequential “snapshots” of animation demonstrating perceptual animacy by Fritz Heider and Marianne Simmel (Scholl & Tremoulet, 2000)	29
Figure 3.2 Genies standing still with their jiggling heads in <i>Princess Mononoke</i> (1997)	33
Figure 3.3 Ripples in Microsoft Word	33
Figure 3.4 Follow-through motion of a tie under the influence of the body movement	34
Figure 3.5 A screenshot of <i>Composition in Blue</i> (1935)	35
Figure 3.6 Use of the jog dial of a VTR	58
Figure 3.7 A screenshot of Sodaconstructor	63
Figure 3.8 A digital shadow theater using Animata	63
Figure 3.9 Snapshots of the iPhone application Spawn	64
Figure 4.1 The diagrammatic relationship of four constitutive parts of the new animation paradigm	70
Figure 4.2 Bird flocking simulated by Craig Reynold’s <i>Boids</i> , modeling different steering behaviors: separation, alignment, and cohesion	74
Figure 4.3 Snapshots of the iPhone application <i>Koi Pond</i>	75
Figure 4.4 A screenshot of an application visualizing the L-system in a tree-like form	76
Figure 4.5 A snapshot of the virtual tree in the website <i>Ecotonoha</i>	77

Figure 4.6 Interacting with <i>Electroplankton</i> using a stylus	80
Figure 4.7 An icon popping up in the dock of Mac OS X	81
Figure 4.8 A snapshot of an execution of <i>Game of Life</i>	83
Figure 4.9 Screenshots in <i>Electroplankton</i>	84
Figure 4.10 Incidental shadow vs. intended shadow	88
Figure 4.11 The interface of the mobile phone N702iS showing computer-generated imagery of water reactive to user action	93
Figure 4.12 The immediate blend of the water-level interface	94
Figure 4.13 The metaphorical blend of the water-level interface	95
Figure 4.14 A mobile Islamic call to prayer application <i>Sun Dial</i>	102
Figure 4.15 Snapshots of <i>SnowDays</i> at Popularfront.com	106
Figure 4.16 BIG SHADOW found in Tokyo	109
Figure 4.17 Digital artwork <i>Text Rain</i>	109
Figure 4.18 The Gorillaz performance in 2006 Brit Awards	110
Figure 4.19 Screenshots of Comeclean.com where confessed messages input by users can be washed away	115
Figure 4.20 A screenshot of <i>Modern Living</i> in Hoogerbrugge.com	116
Figure 4.21 John Whitney's <i>Catalogue</i> (1961)	116
Figure 4.22 The continuum of liveness	117
Figure 6.1 An opening screenshot of Whitney's <i>Arabesque</i>	133
Figure 6.2 The immediate blend in <i>Arabesque</i>	135
Figure 6.3 The metaphorical blend in <i>Arabesque</i>	136
Figure 6.4 A screenshot showing Luxo examining a ball	139
Figure 6.5. The immediate blend in <i>Luxo Jr.</i>	139
Figure 6.6 A screenshot of <i>Pong</i>	142
Figure 6.7 The immediate blend in <i>Pong</i>	144

Figure 6.8 A screenshot of <i>Passage</i> with the player character at the beginning of his life journey	147
Figure 6.9 The immediate blend in <i>Passage</i>	149
Figure 6.10 A metaphorical blend in <i>Passage</i>	151
Figure 6.11 The immediate blend in <i>Koi Pond</i>	156
Figure 6.12 The immediate blend in the genie effect of Mac OS X	161
Figure 6.13 The immediate blend in the water-level interface	164
Figure 6.14 The metaphorical blend in the water-level interface	165
Figure 6.15 Cut-out snowflakes tailor-made by individual users carry their greeting messages evocative of personal memories	168
Figure 6.16 The immediate blend in <i>SnowDays</i>	170
Figure 6.17 The metaphorical blend in <i>SnowDays</i>	171
Figure 6.18 The immediate blend in Comeclean.com	175
Figure 6.19 The metaphorical blend in Comeclean.com	176
Figure 6.20 The immediate blend in traffic lights	180
Figure 6.21 The immediate blend in <i>Text Rain</i>	184
Figure 6.22 The metaphorical blend in <i>Text Rain</i>	185
Figure 6.23 #43 ‘Itch’ in <i>Modern Living</i>	187
Figure 6.24 #61 ‘Drowning’, #54 ‘Jumpy’, & #68 ‘Obedience’ in <i>Modern Living</i>	189
Figure 6.25 #60 ‘New Religion’ & #85 ‘Material Guy’ in <i>Modern Living</i>	189
Figure 6.26 #83 ‘Possessed’ in <i>Modern Living</i>	190
Figure 6.27 The immediate blend in #83 ‘Possessed’	191
Figure 6.28 The metaphorical blend in #83 ‘Possessed’	192
Figure 6.29 #87 ‘Vaudeville’ in <i>Modern Living</i>	193
Figure 6.30 #63 ‘Perfect Day’ in <i>Modern Living</i>	193
Figure 6.31 #71 ‘El Mariachi’ in <i>Modern Living</i>	194

Figure 7.1 Screenshots of <i>Ghost in the Shell</i>	204
Figure 7.2 Screenshots of <i>Princess Mononoke</i>	206
Figure 7.3 Screenshots in the title sequence of <i>North by Northwest</i>	207
Figure 8.1 A diagram showing the associational (link) and oppositional (shift) relationship between the tiles	222
Figure 8.2 A graph showing the edge qualities of the tiles in cinematographic convention: close-up (CU), full shot (FS), wide shot (WS), etc. Please note there exists some tolerance between WS and FS incorporating variability and divergence.	223
Figure 8.3 A perceptual link (conceptual shift) from factory to walkway, with a structural constraint from WS to FS	224
Figure 8.4 A conceptual link (perceptual shift) from amusement park to train track, with a structural constraint from WS to FS	224
Figure 8.5 Rules for transforming personas and some examples	226

LIST OF SYMBOLS AND ABBREVIATIONS

CGI	Computer Generated Imagery or Images
HCI	Human-Computer Interaction
GUI	Graphical User Interface
AI	Artificial Intelligence
VR	Virtual Reality
VTR	Video Tape Recorder
AL	Artificial Life
Mac OS	Macintosh Operating System
MS-DOS	Microsoft Disk Operating System
GVR	Generative Visual Renku
CU	Close-Up
FS	Full Shot
WS	Wide Shot

The text and figures of Chapter 2, Chapter 3, Chapter 4 and Chapter 6, in part, contain material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2009a, 2009b). The dissertation author was a co-author of this paper.

The text and figures of Chapter 8, in part, contain material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2010; Harrell & Chow, 2009). The dissertation author was a co-author of this paper.

The text of Chapter 9, in part, contains material that is a reprint of, or has been submitted for publication in (Chow, 2009). The dissertation author was an author of this paper.

SUMMARY

Animation is popularly thought of as a sequence of still images or cartoons that produce an illusion of movement. However, a broader perspective of animation should encompass the diverse kinds of media artifacts imbued with the illusion of life. In many multimedia artifacts today, computational media algorithmically implement expanded illusions of life, which include images not only moving, but also showing reactions to stimuli (reactive animation), transforming according to their own internal rules (autonomous animation), evolving over a period of time (metamorphic animation), or even generating varying instances subject to user intervention or chance (contingent animation).

Animation in the digital age consists of forms as varied as computer-generated imagery (CGI) in films, motion graphics on interactive multimedia websites, animated contents of video games, graphical interfaces of computer systems, and even digital signage in communal areas. With these forms, the new animation phenomena emerge from entertainment media, functional designs, and expressive works alike, all of which may engage viewers' sensory perceptions, cognitive processes, as well as motor actions.

Hence, the study and creation of animation now requires an interdisciplinary framework, including (1) insights from perceptual psychology and animation studies about animacy, (2) theories of conceptual blending from cognitive science applied to understanding images, (3) notions of embodiment and temporality in phenomenological approaches to human-computer interaction (HCI), and (4) new interpretations of liveness in performance studies accounts of computer-mediated performance. These emergent

ideas jointly characterize the new role of animation in media, and produce new possibilities for more embodied, evocative, and affective forms of generative and interactive animation.

CHAPTER 1

INTRODUCTION

This chapter provides general readers with a background understanding of the many ambiguous meanings of the term “animation” and introduces a new theoretical perspective on the subject. It also states the dissertation’s central thesis question, and introduces the motivating beliefs, aims, and intended outcomes of it.

Animation as a Medium

Animation, to many people, is regarded as a synonym for cartoons, a subcategory of movies or television programs, which primarily target children. Innumerable animated cartoons have been produced over the past hundred years, including many classics from Walt Disney Studios and comparable contenders in North America like Fleischer Studios, Warner Bros. Animation, and others. They gave birth to many cartoon characters, such as Mickey Mouse, Betty Boop, Popeye, Daffy Duck, and somewhat more lately Buzz Lightyear, who constituted the childhood memories of many. The 24-hour rotation of both classic and contemporary animated TV shows on cable television further shows that animation is definitely a major entertainment medium to most Americans (Wells, 2002, pp. 3-6). On the other hand, the viewership of Japanese animation, commonly known as anime, is an antithesis to the North American norm because its target audience is not primarily composed of children, and its themes can be as varied as those of live-action cinema (Napier, 2005, pp. 15-19). Anime, manga (Japanese comics), and video games have garnered so much media time and space in Japan that they have spun off many other

subcultures, like the *cosplay* (costume play) and *otaku*¹ scenes (Lamarre, 2009, p. 144). Many anime fans consider animation as an indispensable part of their lives. In this regard, animation is not just an entertainment form, but also a socio-cultural representation that entails the making of both everyday and extraordinary cultural meanings (Fore, 2007).

Animation as a Technology of Illusion

On the consumer side, animation is a medium. On the producer side, the story is different. To many filmmakers and media producers, animation is the technique of creating an illusion of movement by means of showing successive images at a flickering speed. The technique has been widely applied in film productions for many years to create special effects, such as using the stop-motion technique to make the larger-than-life gorilla in *King Kong* (1933) appear to climb up the Empire State Building. Similarly, by means of the cel-animation technique, the cartoon character Bugs Bunny appears to play basketball with Michael Jordon in *Space Jam* (1996). Lately, with the advances in computer graphics technology, the illusion of life has become more photorealistic. The audiences' eyes are cheated to believe that extinct species of dinosaurs are reborn in *Jurassic Park* (1993), or that Manhattan is destroyed by an immense tsunami in *The Day After Tomorrow* (2004). However, in the eyes of media historians or theorists, no matter how realistic the visual effect on the screen is, it is still composed of a sequence of still images. The essential technique is no different from a set of nineteenth-century bourgeois toys, including the thaumatrope, the phenakistoscope, the zoetrope, and the praxinoscope. All these apparatuses, collectively called animation toys, existed well before film and are believed to have inspired the inventions leading to the cinematic apparatus. Hence, some

¹ According to Lamarre (2009:xvii), the term *otaku* generally refers to “cult fans,” who are “totally into manga, anime, video games, and a range of related merchandise and events.”

theorists (Cholodenko, 1991b, pp. 9-10; Manovich, 2001, p. 302) see animation as the ancestor of the cinema, rather than a part of it, in technological terms.

Animation as an Art Form

Regarding animation as just a set of technologies induces annoyance in many artists who have been fascinated by, or obsessed with, the specific medium for quite some time. Instead of just a way to create illusion, many of them see animation as a singular art form enabling them to convey messages in unprecedented and expressive ways as they navigate through a continuum between mimesis and abstraction (Furniss, 1998, pp. 5-6). The unique medium benefits from live-action film the poetics of narratives, and shares the aesthetics of ambiguity with the visual arts. Hence, it has always been one of major playgrounds of the formal avant-garde. From Walter Ruttmann's *Opus I* (1921) and Hans Richter's *Rhythmus 21* (1921), to the later works of Oskar Fischinger (e.g., *Composition in Blue*, 1935), Len Lye (e.g., *A Color Box*, 1935), Norman McLaren (e.g., *Hen Hop*, 1942), experimental abstract films from that era involved animated components substantially enough that they can even be classified as animation films. Parallel to these radical attempts, some others have explored the nuances of narrative mode and the physicality of the medium. Chuck Jones (e.g., *Duck Amuck*, 1951), Osamu Tezuka (e.g., *Broken Down Film*, 1985), and lately Daniel Greaves (e.g., *Manipulation*, 1991) tell stories in their exemplars without forgetting to remind the audience of the specificity and materiality of the medium, such as the camera frame, film celluloid, paper, ink, and paint. Meanwhile, there are numerous canonical works that depict poignant, evocative, and metaphorical stories, including Jiri Trnka's *The Hand* (1965), Caroline Leaf's *The Street* (1976), Yuriy Norshteyn's *Tale of Tales* (1979), Frédéric Back's *The Man who Planted the Trees* (1987), and Michael Dudok de Wit's *Father and Daughter* (2000). As today's computers and software applications have become accessible, independent animators can

create works that are totally self-driven. Now animation is an art form accessible to the independent auteur like painting and sculpture.

In summary, animation is many things. It can be a kind of entertainment. It can also be a form of representation carrying socio-cultural meanings. On the one hand, it refers to a set of technologies creating the illusion of movement. On the other hand, it is an art form negotiating the tension between reality and imagination. Animation might refer to different concepts for different people within different contexts. Some might also say that a piece of animation art can simultaneously apply certain animation techniques to create illusion, be socially critical, and be entertaining. In other words, all these different interpretations of animation in the meantime overlap in some areas. As I have pointed out at the beginning of this chapter, the term “animation” is ambiguous.

Animation in the Digital Age

Just like other media or art forms, animation has undergone a digital revolution over the past decades. First, we have witnessed animation proliferating in a wide array of entertainment media, from animated cartoons, motion titles, music videos, video game contents, to interactive multimedia websites, outdoor electronic displays or projections. Second, major production technologies have advanced from the traditional hand drawings on cels, articulated three-dimensional puppets or two-dimensional cutout silhouettes, or deployment of animatronic figures, to computer-generated characters and computer-simulated visual effects. Third, more and more works of art or performance incorporate animation into their video or projection content. Most revolutionary of all, many interfaces of personal computers or hand-held devices have incorporated animated graphical elements. Although these insertions are seemingly for the sake of adornment, they are believed to enhance agency of the systems and constitute continuity of the human-computer dialogues (Suchman, 2007, pp. 210-214). In short, animation is

becoming perceptually pervasive today. We can see animation everywhere in the digitally mediated environments we inhabit.

In light of this current animation landscape, my overall questions are: during the current digital revolution, do these convoluted animation phenomena suggest new approaches to the analysis and design of new digital artifacts? Can established theories and practices in one area, such as research in user-interface design, give new insights to another, such as the study of animation media, and vice versa? I believe that the creation and study of the emergent cross-domain animation phenomena in digital media would require an interdisciplinary approach.

Animation as a Phenomenon of Liveliness

According to the origin of the term, the literal meaning of “animation” in Latin is “instilling with life.” “To animate” is to give life to something. Since what or who gives life is always unknown, the act of endowing with life inevitably leads to a spiritual end, such as for alchemists in medieval Europe or the Taoist’s in ancient China. To manifest the idea of life, some writers use the word figuratively to describe something that “looks” lively but is not necessarily alive, for example, “animation of the machine” (Turkle, 1984, p. 27). In this sense, life is a kind of everyday experience and animation is an illusion of that experience. In the traditional view of animation, the illusion is mainly projected by means of movement. I argue to expand the meaning of “illusion of life.” This illusion hinges not solely on movement, but on the perceptual phenomenon of liveliness, which is reminiscent of our embodied everyday experiences with life, such as reaction, metamorphosis, growth, and so on. It is my belief that this interpretation gives us a useful theoretical lens to juxtapose and analyze animation phenomena of different domains in comparable vocabularies. When an animated cartoon cat walks like an elegant lady, when two squares in an abstract film spin like a pair of dancers, when computer interfaces present windows moving playfully and responding to user clicks by twisting

and minimizing, when game characters bounce, squash, and stretch while heroically exploring their worlds, the phenomenon of liveliness is the soul of all animations.

This phenomenon-oriented approach to animation tends to include everything around us in the world. A counter-question is: is it theoretically fruitful to have an all-embracing concept of animation? My aim is not to invent an all-encompassing theory of animation, but rather to mobilize the interlinking, or to destabilize the borders, between different areas of knowledge pertaining to animation. It is not my intention to establish overarching knowledge of such areas as human-computer interaction, artificial life, artificial intelligence, digital art, or digital performance. I only believe that each of these areas that manifest the illusion of life would help delineate the essence of animation in the digital age. My ultimate focus is still on how images, with the aid of computers, can project an expanded illusion of life.

Hence, this dissertation is intended to

- theoretically establish a new perspective of animation in today's digital age as a perceptual phenomenon (Chapter 2);
- extend theories of perception, cognition, and bodily interaction for analysis based in this new perspective on animation (Chapter 3 and Chapter 4);
- examine a corpus of animated artifacts exemplifying the new animation paradigm, and identify a set of key variables defining the primary concerns for designers and developers (Chapter 5 and Chapter 6);
- suggest a novel framework for designing more embodied, evocative, and affective forms of generative and interactive animation (Chapter 7 and Chapter 8);
- investigate potential contributions to new principles of interface design, and new insights for principles of animation at large (Chapter 7).

CHAPTER 2

PROBLEM FORMULATION

This chapter starts with characterizing the emerging animation phenomena I have observed from today's digital media and have introduced in the previous chapter, followed by questions I raise to challenge the inadequacies of existing approaches to the study and creation of animation. I then propose a "movement" away from analytically focusing on any single medium to studying a more embodied experience of the illusion of life with lessons informed from several peculiar mechanical artifacts, which give rise to the idea of generative and interactive animation.

New Animation Phenomena in the Digital Age

As mentioned in the last chapter, with recent advances in multimedia technologies, animation has become increasingly pervasive. Apart from the cinema and television, today we can see animation in many digitally mediated environments, such as the graphical user interface (GUI) of many operating systems, interactive multimedia websites on the Internet, entertainment or lifestyle applications on hand-held mobile devices, and many electronic displays or projections in outdoor communal areas. These multimedia artifacts consist of "animated" (in the sense of "bringing to life") images, which are not only "moving," but also presenting an expanded illusion of life, such as showing reactions to stimuli (reactivity), transforming according to their internal rules (autonomy), evolving over a period of time (generativity or metamorphosis), or even producing divergent outcomes subject to chance or intervention (contingency). Examples of reactivity include a graphic icon in Macintosh OS X bouncing restlessly in response to a user click, and a mouse pointer shedding glitter like fairy dust trailing Disney's version of Tinker Bell. Autonomous animation can be seen when a cartoon dog runs on a

computer screen during a file transfer, or when a little green man walks faster and faster in a traffic light on counting down to the red signal (Figure 2.1). A virtual tree growing on a website based on user-submitted content demonstrates both metamorphosis and contingency. Despite the ubiquity, however, such animation phenomena are typically not considered in the study of animation, but instead are regarded as visual effects in user-interface design, computer game development, and related fields. I believe that the idea of animation should be expanded to encompass this wide array of phenomena of liveliness enabled by emerging technologies of digital media, because they altogether project a more pervasive impression of life than traditional filmic animation.



Figure 2.1 An animated “running” green man on a traffic light in Taipei

Theorizing the New Animation Phenomena

Considering the new animation phenomena in digital media based on existing animation theories, which primarily centralize traditional filmic animation, inevitably brings about challenging theoretical inquiries. Viewing experience of the new animation phenomena differs from that of filmic animation in that a viewer of the former is not just a passive audience but also a participant. Initially, the viewer perceives movements in

digital media the same way as in animation films and interprets the images subject to her or his past movie or television viewing experience. Yet, the viewing context in digital media may differ considerably. In addition to perceiving the images on a screen, the viewer is invited or required to “fiddle” with the artifacts, such as configuring a GUI on a mobile phone, controlling a game character, or interactively operating a multimedia authoring software application. In short, when the viewer sees the animation, she or he also interacts with the artifact. This scenario triggers us rethinking many traditional perspectives on animation. For instance, the genie effect of Macintosh OS X is perceived differently from the genie in *Aladdin* (1992) (Figure 2.2). The little green man walking and running on traffic lights does not work the same way as the incredible Hulk in both functional and perceptual terms. Textual analysis² in terms of semiotics, genre, or narrative that formerly applied to animation films or television shows might not be directly relevant to the new animation phenomena. Regarding the functionality of these artifacts, their goals and objectives might as well be completely unrelated to conventional animations. Before looking into these issues, we need to start with an overview of some existing approaches to animation, from both the humanities and computing disciplines.

² The meaning of the term “text” here follows Roland Barthes, who sees visual media such as photography, film, television, or painting based on codes, like verbal languages.



Figure 2.2 Two genies, one in a cartoon animation, another in Macintosh OS X

Existing Approaches to Animation

In this section, I present a brief survey of approaches to studying animation from film-based and computer graphics-based perspectives. I then point out the inadequacies of these approaches by highlighting their over-reliance on considering the medium-specific images in animation as the central object in their analyses. These medium-centric approaches inevitably fail to cope with the new paradigm of digital animation driven by emerging technologies across multiple media.

Film-Based Approaches

In humanities disciplines, animation is generally seen as a marginal type of film, often referred to as an “animation film.” Scholars have predominantly drawn on two approaches to animation research, namely contextual and textual analysis (Furniss, 1998, pp. 7-12). The former looks at production contexts, including the historical, industrial, technological, and economic situations in which individual works can be understood. Some studies in this approach also review the national or cultural aspects of some animation films within a film genre. For example, anime is lately an emerging genre, which can be studied as a national cinema of Japan with respect to its post-WWII history

of economy, just happening to be in the animation category. Susan Napier's research in anime belongs to this type of studies, uncovering vital links in both global and local contexts (Napier, 2005). The latter approach, usually more theoretical and ahistorical than the former, draws attention to the canonical texts of the specific medium. It entails conducting close readings of works and applying theoretical models to perform analyses of the aesthetics, semiotics, or narratives. For instance, Osamu Tezuka's *Broken Down Film* (1985) can be contrasted with Chuck Jones' *Duck Amuck* (1953) and Daniel Greaves' *Manipulation* (1991) in terms of the self-reflexivity of cel animation embedded in the narratives. Paul Wells's comprehensive typology of narrative strategies employed in a set of canonical animation films demonstrates how this method is applied to analyzing the distinctive medium, suggesting a wealth of possibilities for animators and filmmakers to create their new works (Wells, 1998, pp. 68-126). While this cursory overview of animation studies does not encompass the entirety of film-based approaches to animation, it captures at least a sketch of the prevalent and conventional approaches within the field.

Although research in film-based animation has contributed an immense and powerful corpus for the analysis of works creating the illusion of life with sequential images, these approaches mainly concentrate on the perception and meaning of visual images without addressing the embodied engagement with many multimedia artifacts that are also imbued with phenomena of liveliness. The prevalent view of animation as only a type of film is too medium-centric in that it falls short of explaining the new animation phenomena that have clearly emerged in various media. When computer interfaces contain icons that jump restlessly in response to user clicks, and when game characters squash and stretch in the game worlds according to players' control, designers and developers working in cross-media fields today require new insights and principles to design and produce digital media artifacts of this new genre. We need to import new areas of knowledge beyond the film medium.

Computer Graphics-Based Approaches

On the other hand, in computing disciplines the research agenda of animation has primarily been about the digital synthesis of a sequence of images. Early research areas included hierarchical modeling and animation in primitive graphic forms, and development of efficient standards and encoding for multimedia content. These concerns have been expanded by later studies to develop works including computer representation of physical objects and materials, photorealistic rendering (and more recently stylized rendering, such as cel-shading) techniques, simulation of physical phenomena like particle systems and spring dynamics, developing algorithmic approaches to generate organic motion such as flocking (e.g., Craig Reynolds's *Boids*) or other self-evolving patterns like cellular automata (e.g., John Conway's *Game of Life*), implementing artificial intelligence (AI) programs to create animated behaviors³, and much more. Initial application domains were primarily technical, scientific, medical, architectural, and recently cinematic, yet the marvelous illusions generated by computers have spread to other communal and personal entertainment platforms, including television, digital signs and billboards, notebook computers, and hand-held devices.

Mainstream computing disciplines see animation as a part of computer graphics. The term computer animation is narrowly understood as only computer-generated imagery (CGI), which is also a kind of medium-specific images. The promising capability of generative algorithms supporting interactive construction of polymorphic instances of animation has been downplayed. However, the animation phenomena in GUIs and other digital environments remind us that a new paradigm of animation is

³ Though work in this vein such as in Bates (1993, 1994) often focuses on creating digital characters with a sense of believable "aliveness," which often highlights creating seemingly intelligent and emotional human-like responses more than the type of perceived dynamic phenomena that is the focus of my concern with "liveliness," see Chapter 7 for more discussion.

emerging on the computer platform. This generative and interactive form of animation has not been addressed directly by current research in human-computer interaction (HCI), because its concerns are mainly usability- or productivity-oriented. Although a few other researchers like Joseph Bates (1993, 1994), Michael Mateas (2004), and Ken Perlin have developed works that do consider computer expressiveness as part of their concerns, they have not explicitly pointed out the role of animation in computational media. We need an interdisciplinary approach that brings consideration of expressiveness and entertainment value of animation to the design and creation of computer artifacts.

The above approaches to animation are not only medium-centric, but are also discipline-centric. On the humanities side, major studies of animation emphasize the aesthetic, expressive, affective, and social- or even political-critical values of the artifacts. Scholars in these areas are interested in how animation makes meaning. On the computing side, researchers focus on the practical, informational, or entertainment potentials of the related technologies. They incline to explore how people and society can benefit from a state of the art. The two disciplines contrast dramatically in their mind-sets.

Definitions of Animation

Aside from being colonized by either Film Studies or Computer Science, animation has also been isolated as a separate area of inquiry as some animation theorists are interested in marking a new territory. At the outset, the majority of these theorists define animation by its creation processes and industrial practices. This majority unanimously arrives at a definition that hinges on the frame-by-frame manipulation of images (Ward, 2000). On the other hand, some define animation in terms of styles. Maureen Furniss outlines a continuum of animation in relation to live-action imagery. She uses “abstraction” and “mimesis” to represent the extremes of the continuum, which aims to encompass all different film styles and approaches (Furniss, 1998, pp. 5-6).

Similarly, Paul Wells proposes a relational array between “experimental” and “orthodox,” with a remark that animation should challenge the regime of live-action films (Wells, 1998, p. 28).

In contrast, the Canadian animator Norman McLaren (1950) came up with what has proved to be the most influential definition of animation, which is more about essence than the process: “animation is not the art of drawings that move but the art of movement that are drawn” (Furniss, 1998, p. 5). His emphasis on movement rather than drawings echoes Gilles Deleuze’s view: “animation is not the transformation of images into one another, but the images being formed and dissolving through movement” (Schaffer, 2007). That is to say, animation is not formed by images before movement but rather after it. That means movement precedes images in animation, for there are no animated images without movement. Both McLaren’s and Deleuze’s views seem to imply that movement constitutes animated images. In other words, movement is the essence of animation. Their ideas, prioritizing movement over images or drawings, mark a major step toward the illusion of life. However, they still position animation in a particular medium. Considering only the cinematic medium is inadequate in addressing the emerging phenomena pertaining to animation across multiple media. As I have mentioned in the last chapter, we need a broader perspective of animation to investigate the new cross-media and cross-discipline animation phenomena.

Beyond Single Medium

Hence, I call for a “movement” away from centralizing any single medium to a phenomenon-oriented approach to the study of animation. This movement brings into consideration human perception, cognition, and bodily interaction when dealing with animated artifacts, addressing the aforementioned discipline- and medium-centric problems. On the one hand, the new perspective mobilizes interplay of knowledge between different areas, including perceptual psychology, animation studies, visual

studies, cognitive science, HCI, and even performance studies. It echoes with, and actually extends, Paul Ward's suggestion for a discursive study of multiple areas of knowledge about animation (Ward, 2003). On the other hand, the movement shifts our focus from medium-specific images to diverse kinds of media artifacts imbued with an illusion of life, spanning a range from the nineteenth-century bourgeois toys to computers in the late twentieth century. The close relationship between animation and many early experimental artifacts creating moving images has been respectively marked by Alan Cholodenko (Cholodenko, 1991a, 2007), Lev Manovich (Manovich, 2001), and others. To me, Cholodenko's latest view is exceptional in that he draws attention to the animatic apparatus in addition to the animated image (Cholodenko, 2007). Cholodenko thinks that the "animatic apparatus," which he refers to the mechanical apparatus that generates moving pictures, such as the phenakistoscope or the zoetrope, manifests a double definition of animation: simultaneously meaning "endowing with life" and "endowing with movement." One could interpret that it animistically gives life to inanimate images through mechanical movements like spinning, flipping, and illuminating. In this regard, I would add that animation should not only refer to visual imagery and the animatic apparatus that generates it, it also includes the operation of the apparatus, which performs the embodied and material realization of animation. Early animatic apparatuses seem on the surface level to involve only visual perception. Yet in practice, a certain degree of motor action on the viewer's part is required, from the simplest action of peeping in, to more engaging physical operation to "generate" and "maintain" the moving images. Hence, it should be recognized that sensorimotor interaction is essential to the operation and reception of animation.

Cholodenko calls his readers' attentions to the animatic apparatus, including accounts of animation toys. These artifacts are generally regarded as ancestors of animation in Euro-American study of the arts. In fact, many mechanical artifacts from other parts of the world also manifest the human pursuits for creating an illusion of life

and movement. Most importantly, they all mobilize a connection between motor action and sensory feedback on the viewer's side. They include shadow puppets, humanoid automata, *karakuri* (a kind of Japanese mechanical robot developed in the Edo period), pacing-horse lamps (a kind of traditional Chinese lanterns), and Chinese handscroll paintings.

Shadow puppets are brought to life under the motor control of puppeteers. While moving a rod hooked up the limb of a shadow puppet, a puppeteer is simultaneously a performer as well as an audience of her or his own show, because the puppeteer also gazes at the animated images on the screen, just from the other side of it compared with the theatrical audience. The puppeteer's visual and motor systems are wired in the movements of the puppet (Figure 2.3).

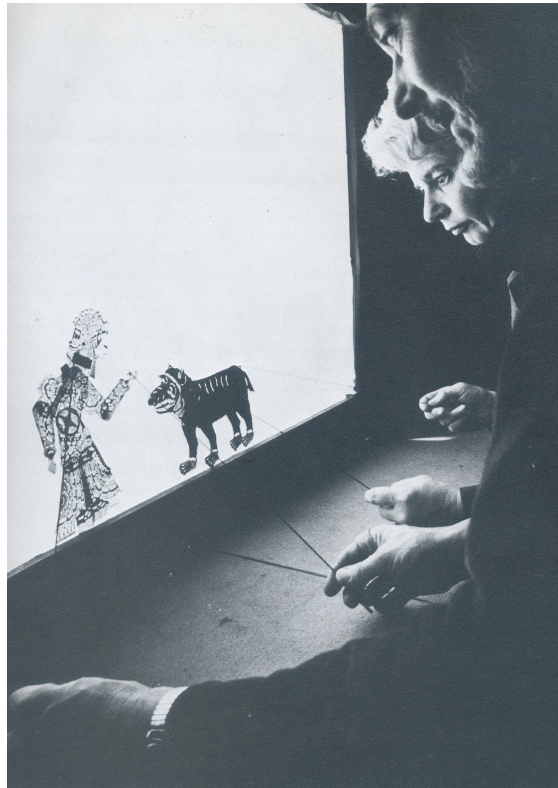


Figure 2.3 Puppeteers being simultaneously performers and audiences of their own show

Automata are “programmed” by humans through mechanical means to “demonstrate” self-movement, creating an illusion of autonomy. Their control is largely

limited to the basic on-and-off operation. Yet, some of the humanoid automata, such as the Japanese tea-serving karakuri (Figure 2.4), can demonstrate an embodied and situated interaction with the audience. When the audience takes the cup of tea from the karakuri's hands, it stops pacing and waits patiently. Once the teacup is returned, it bows courteously and turns back. The action and reaction loop is embedded in the ritual of Japanese tea ceremony. Although puppets and automata are peripheral to human actors in theatrical performance, they manifest the human pursuit for creating the illusion of life.



Figure 2.4 A tea-serving karakuri

Meanwhile, the Chinese pacing-horse lamp (Figure 2.5) functions more like a peculiar mix of mechanized zoetrope and shadow play. The turning of the drum inside the lantern is driven by thermodynamics. Although the mechanical process is different from that in automata, the human operation is similar. The viewer is always able to adjust the turning speed for an optimal viewing experience.

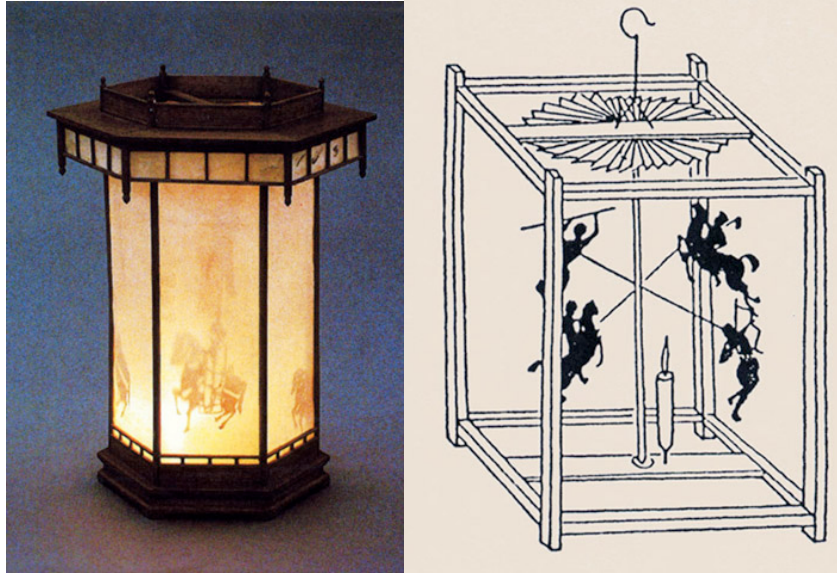


Figure 2.5 Figures on a pacing-horse lamp spinning automatically

Besides, many Chinese handscroll paintings entail even more viewer-driven interaction. The handscroll, a distinguished format in Chinese painting, is an exceedingly long horizontal scroll. When not in use, it is rolled up. On viewing, the beholder unrolls part of it, approximately an arm-length at a time. This length roughly defines the size of the viewing frame, which slides over the whole scroll, like a horizontal pan of background in cel animation. This paper panning is coordinated by both hands of the viewer and loops back to her or his eyes. When the viewer's field of sight approaches the edge of the frame, the viewer would gently unroll that side of the scroll and roll up the other end. The motor action of the hands and the visual perception jointly produce an apparent sense of controlling a horizontal camera pan. Although the picture is undoubtedly static, the peculiar artifact does give the beholder an experience of moving camera view in a landscape or even a narrative world (Figure 2.6).



Figure 2.6 A beholder's hands and eyes coordinating the panning of a hand-scroll painting (Wu, 1996, p. 58)

The causal relationship is intricate between the illusion of life created and the motor-sensory connection embedded in these artifacts. Apparently, the illusion is triggered and maintained by motor control, and then perceived by the user. Conversely, the created illusion amazes the audience and incites further bodily engagement with the artifacts. This engaging feedback loop characterizes an embodied experience of an expanded illusion of life that involves not only sensory perception but also motor action.

It is this feedback loop that is a major characteristic of animation on computers today in which algorithms support interactive multiple generation of the expanded illusion of life. Similar interplay between illusion and engagement in the above mechanical artifacts is also active in many digital media artifacts. We are presented with digital puppets (e.g., avatars in computer games or virtual worlds), computational automata (e.g., screensavers of computer operating systems), and interactively scrollable pictures (e.g., Apple Quicktime VR technology). These terms are not metaphorical, but rather are ontological, because computers actually display reactive, autonomous, evolving, and contingent images, just like percepts that shadow puppets, *karakuri*, pacing-horse lamps, and the like, convey to their audiences. When we walk across the

road “with” the little green man on a walk signal, when we call in a genie-like application on a computer, when we move the mouse pointer with trailing fairy dust like a magic wand, when we tilt a hand-held gadget to pour water “out of” the screen, animation alongside the computer provide us with a more embodied experience of an expanded illusion of life than the traditional film form of animation.

I have been calling this the “new animation paradigm,” “generative and interactive animation,” or most simply, interactive generative animation. This type of animation is generative, because works of this kind are able to present different versions rendered possible by computational generation. As Philip Galanter puts it, generative art is a method in which an artist off-loads “moment-to-moment intuitive decisions to an external *autonomous system*” (Galanter, 2006, my italics). By autonomous system, he meant a system generating time-sensitive instances. Hence, the multiple instances of the new animation works are dynamically generated by systems. Meanwhile, works of this new paradigm are also interactive, because they are often open to user input and intervention. Simon Penny asserts that interactive art should involve machine *systems* to produce *reactive instances* based on sensory input (Penny, 1996). In other words, systems generate different versions of output in response to different user input on different occasions. It follows that new animation works are made interactive with the support of systems too.

Both Galanter’s generative art and Penny’s interactive art include systems and instances, which give rise to major phenomena of liveliness I have reiterated, namely the autonomous and the reactive. Moreover, as I shall show in later chapters, other phenomena like contingency would also emerge as a result. All in all, generativity and interactivity are intricately conjoined in the new animation paradigm, creating an expanded illusion of life in computational systems. In the next chapter, I shall underscore several established areas of knowledge that are useful for us to investigate the characteristics of this new type of animation from multiple perspectives.

The text and figures of this chapter, in part, contain material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2009a, 2009b). The dissertation author was a co-author of this paper.

CHAPTER 3

THEORETICAL FRAMEWORK

In digital media, the new animation paradigm gives audiences an embodied experience of an expanded illusion of life, which includes not only movement, but also various other phenomena of liveliness, such as reactions to stimuli, rule-based autonomous behaviors, or even enduring and contingent transformations like metamorphosis or growth. These emerging animation phenomena are becoming more pervasive, but also increasingly convoluted, in our digitally mediated everyday life. What surround us are apparently embellishments, but I believe they are in fact representing an enduring human concern: to create an illusion of life. We require a more robust framework to uncover the underpinnings for notion of the new paradigm of animation. To start analyzing the new animation phenomena, I delineate the necessary conditions as follows:

- There must be a creator who designs, plans, or programs an illusion of life.
- There must be an *observer* who perceives the illusion.
- There must be material manifestation including *imagery* of which people make meaning.
- There must be *bodily engagement* with the material artifact.
- In digital environments, the creator, the observer, and the artifact co-enact a new form of live *performance*.

These conditions mark the human perception, cognition, and bodily interaction in relation to animated artifacts in general. Based on these components, I propose four areas of knowledge as the points of departure:

First, I look at the new animation phenomena from the observer's perspective (The Observer). I review thoughts from animation studies and perceptual psychology

regarding how we “perceive” certain dynamic phenomena that are reminiscent of our everyday experiences with life. In short, I invoke insights about the observer’s, as well as the creator’s, perceptions of animacy. References include (Arnheim, 1974; Cholodenko, 1991a, 2007; Dennett, 1987; Scholl & Tremoulet, 2000).

Second, I examine the interpretation of an image (The Image). I draw upon theories from cognitive semantics and recent findings from neuroscience to discuss how we “read” and “sense” animated images and how these images activate our imaginations. References include (Fauconnier, 2001; Fauconnier & Turner, 2002; Glenberg, 1997; Hiraga, 2005; Hutchins, 2005; Lakoff & Johnson, 2003; Lakoff & Turner, 1989; Turner, 1996).

Third, meaning emerges from the new animation phenomena not just through image viewing but also by taking bodily action (The Body). Starting with introduction of feedback mechanisms from cybernetics and phenomenological approaches to human-computer interaction, I look into how our bodies engage or even “interact” with these animated artifacts. During interactive engagement, meaning is embodied in our bodily action. References include (Dourish, 2001; Gibson, 1986; Merleau-Ponty, 1962; Moran, 2000; Shneiderman & Maes, 1997; Wiener, 1961)

Fourth, the new animation phenomena can be regarded as a mediated form of performance (The Performer). Based on arguments in performance studies about the impact of media technologies, I investigate how real-time control and rendering technologies entangle the notions of performance and animation, and influence the relationship between the performer and the observer, bringing contingencies of life to animation. References include (Auslander, 1999; Dixon, 2007; Kleist, 1810; Lewis, 2000; Reiniger, 1970; Tillis, 2001).

Finally, each of these four areas gives rise to new theoretical ideas about animation articulated in the next chapter.

The Observer: Perception of Animacy

This section addresses the general question: what dynamic phenomena make things “look lively” to an observer. These phenomena of liveliness are reminiscent of our everyday experiences with life and make us aware of life and living things in our mundane world. Yet, this awareness of life and living things should not be directly equated to aliveness, because the former is only a perception, which alone cannot confirm if something is alive. There is subtle difference between “looking lively” and “looking alive” since the latter also hinges on our common knowledge (e.g., someone might be unconscious, but still alive). In other words, I am not aiming to inquire about the metaphysic of life, but instead to characterize the illusion of life. Toward this end, I first differentiate the nuances between the concepts of aliveness and liveliness. I draw upon perceptual psychology and animation studies to delineate phenomena of liveliness of different kinds, namely prominent behaviors and emergent patterns. The former concentrates observers’ attentions on particular progressive actions and behaviors, while the latter dilute observers’ attentions over a collective transforming whole. In order to better illustrate the idea for readers of different backgrounds, I provide the following chart of vocabularies in relation to movement, action, visual components, happening, animation, and film at large as examples, with the concentrated attention on the left whereas the diluted on the right:

Moving in space	Transforming in shape
Chasing	Wandering
Boxing	Dancing
A tug of war	A flock of birds
Dropping a knife	Raining in the environment
Animated figures	Evolving background
Primary animation	Secondary animation

Character animation	Natural phenomena animation
Narrative film	Abstract (nonrepresentational) film

Meanwhile, this new perspective on liveliness also attempts to destabilize or problematize the dichotomy of the animistic and mechanistic views applying to animation, which corresponds to the two conventional animation styles, namely organic versus mechanical motion. At the end of this section, I bring up a message that the new typology of liveliness would lead us to better motion design than the old classification into organic and mechanical movements.

The Animistic View of Life

The connection between the perception of dynamic phenomena and the belief in life and soul is a historical thought in many cultures, such as the alchemists targeting immortality in the West, the notion of Tao in China, and Shinto in Japan. The idea can be collectively called the animistic view of life, in accordance with Alan Cholodenko's essays on the illusion of life (Cholodenko, 1991a, p. 16, 2007, p. 487), as well as Paul Wells's book (Wells, 1998, p. 32). People with this belief might attribute life and soul to many dynamic phenomena, not only humans talking and animals walking, but also plants growing and even natural phenomena, such as leaves waving in the breeze, water running down the river, lightning striking a tree, and the like. They believe in the spiritual connection between life and movement. This view has been denigrated as unscientific and "primitive," often grossly mischaracterized by anthropological texts – that is not the type of argument being made here. In this chapter, the term "animistic" specifically refers to a kind of subjective association of dynamic phenomena with living things. As Mark Turner puts it, we tend to compare an actor moving a physical object to the movement of our body, believing that there is a soul responsible for the bodily movement (Turner, 1996, p. 21). Likewise, one might on a preconscious level believe that falling leaves and

floating clouds all have their own internal impulses and souls. This perspective, as applied to performing art or animation, partly explains the psychological effect when we view a marionette walking on the theatrical stage, a tea-serving karakuri (a mechanical robot) nodding its head, or a computer-generated desktop lamp jumping on the cinema screen. We are inclined to “suspend our disbelief” and see them as alive. In other words, when we see things moving, no matter for a hand drawing, a puppet, or an everyday object, we tend to suppose they are animated (i.e., brought to life), resulting in what we call “animation.”

Piaget on Aliveness

The animistic view also has an empirical grounding. Since 1920s, the Swiss philosopher Jean Piaget started to study children’s perspective of life and soul by conducting interviews. He later concluded in his theory of child development that children are having different criteria of aliveness in the course of their growth. At the very first stage, they consider nearly everything as alive. At the second stage, movement differentiates the animate from the inanimate. At the third stage, they think something that generates its own movement is alive, while something moved by others is not (Arnheim, 1974, p. 400; Turkle, 1984, p. 30). His findings explain why an infant child is engaged with the mobile above the crib, and why kids love the anthropomorphic sun and moon. This theory about child psychology at different ages seems largely compatible with the animistic view that movement instills life. However, one should not hastily think that this view only applies to infants or children, just as some people see animation as cartoons only for kids. In fact, this psychological view at earlier ages is still traceable in our immediate perception after growing up.

Arnheim on Liveliness

The perceptual psychologist Rudolf Arnheim has carefully articulated the nuances between aliveness and liveliness. For Arnheim, the former is a psychological belief, while the latter is a perceptual quality. On the basis of perception and observation alone, we may find something very lively but may not think it is alive at second thought. Arnheim asserts that what counts as liveliness is not whether there is really a mind or soul, but rather the level of complexity of the observed behavior. He delineates different degrees of liveliness according to different levels of complexity, from simple movement to complex behavior (Arnheim, 1974, p. 401), as shown in the following list:

1. Something that moves is livelier than those do not.
2. Movement involving internal change (i.e., change in shape) is at a higher level of complexity than rigid object displacement.
3. The thing moving by its own force (i.e., self-initiated movement) is higher in degree of liveliness than that physically moved by others.
4. Those self-movements initiated by internal impulses are livelier than those driven by external forces.

The first three levels cover many natural phenomena and everyday physical interactions between actors and objects, such as a leaf falling, a balloon deflating, or hitting a tennis ball, whereas the last level corresponds to numerous other organic behaviors like chasing, escaping, searching, avoiding, hesitating, and so on. In Arnheim's sense, an observer sees a movement with a higher degree of liveliness, and so more "animate," if the visible "interplay between force and counterforce," in terms of changes in course, dominance, or others, is becoming more complex (Arnheim, 1974, pp. 402-403).

From Interaction to Intention

Many other researchers in perceptual psychology have also closely investigated how motion influences the perception of animacy (Scholl & Tremoulet, 2000). Fritz

Heider and Marianne Simmel are among the first of those creating short films of geometric figures in motion (Figure 3.1) and asking observers for their immediate feedback. They found that observers saw those figures with personalities and emotions just because of temporal contiguity and spatial proximity in the films. Their findings have called attention to how audiences respond to various qualities of motion, like speed, timing, direction, trajectory, and so on. For example, Winand Dittrich and Stephen Lea presented observers with a set of randomly moving letters among which there is an exceptional one supposed to be intentionally chasing other, like a predator or an attentive follower. They found that for one to be detected as particularly animate, apart from visible *interaction* in terms of movement, the impression of *intentionality* produced by the movement is also equally important. While the perception of interaction depends on what Arnheim calls the visible interplay between the chaser and the chased, the perception of intentionality hinges on only the trajectory of the chaser. A chaser that moves prominently to its goal is more easily detected as animate. Steward also conducted similar experiments by presenting a set of motion that seemingly violates Newtonian laws, that means, showing not interaction between forces, but rather unexpected motion. The upshot is that an object abruptly changing direction to avoid an obstacle or suddenly accelerating toward a goal is perceived to be more animate (Dittrich & Lea, 1994). Lately, Patrice Tremoulet and Jacob Feldman demonstrated that just change in direction and change in speed, even without any goal or obstacle, can produce an impression of animacy. In short, prominent and solo movement against laws of physics also seems lively.

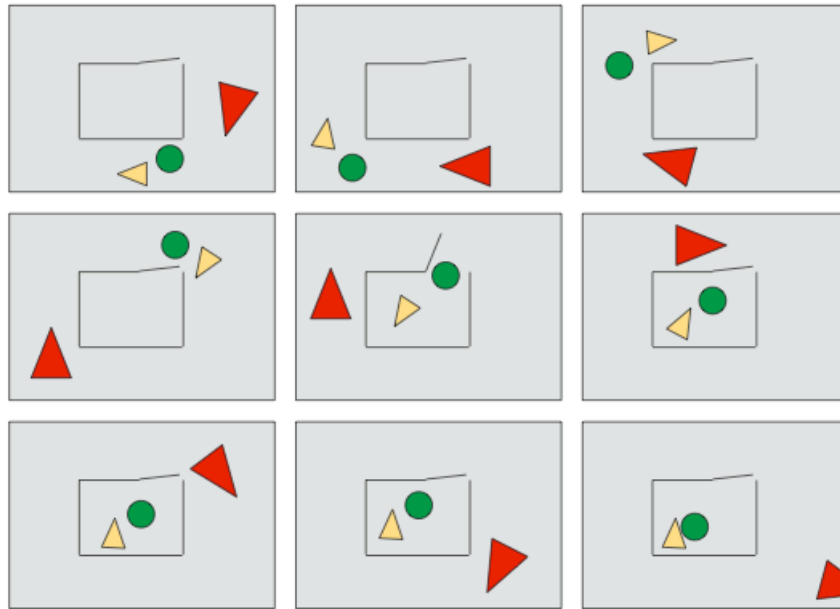


Figure 3.1 Sequential “snapshots” of animation demonstrating perceptual animacy by Fritz Heider and Marianne Simmel (Scholl & Tremoulet, 2000)

Dennett on Intentional Behaviors

Results from psychological experiments seem to be in line with Daniel C. Dennett’s theory of the intentional stance, which differentiates organic behaviors from physical phenomena. The theory describes three levels of abstraction an observer adopts when viewing the behavior of an object. At the most concrete level (the physical stance), an observer uses his knowledge of the laws of physics to predict the outcome of a behavior, for example, a flying ball going to break a window. At the next level (the design stance), an observer assumes an engineered or biological system governing the outcome and ignores the mechanics within the system. For instance, one can predict an alarm clock is about to ring when the set time comes. At the most abstract level (the intentional stance), an observer treats an object having intention toward a goal (Dennett, 1987, pp. 16-17). When we see several ponies fleeing the forest, we guess they are escaping from a danger. The intentional stance not only applies to plants or artifacts (Dennett, 1987, pp. 16-22), but also reflects in our viewing of animation. When we watch

a pendulum, we take the physical stance to predict the swing. In case the pendulum suddenly swings past the highest point and spins around, we would feel that it gains “celestial” internal energy that acts against the natural force, which supposes to keep it on the right track. The unanticipated movement would incite us speculating the intent behind the dynamic and switching to the intentional stance, by which we are not asking whether the pendulum has an inner soul, but rather predicting what the next action is.

To some researchers exploring expressive potential of computation such as George Lewis, Fox Harrell, and Jichen Zhu, this subjective attribution of intention to unanticipated and independent behaviors is even a key to designing seemingly intelligent systems (Harrell & Zhu, 2009; Zhu & Harrell, 2009). Harrell and Zhu think that the sense of intentionality in these systems not only depends on computational complexity, but also results from the interplay of user agency (affecting the system according to its affordances) and system agency (the user’s interpretation of the system’s behavior as self-motivated, responsive, and intentional) in social and cultural situations. I add that temporal context is also crucial, because Dennett’s the intentional stance is time-dependent. Dennett believes that humans would “choose” different levels of stance when viewing the behavior of an object. This act of choosing may not be a conscious mental activity, but rather takes place at the preconscious cognitive level (Dennett, 1987, p. 22), especially when the behavior happens in a time-sensitive situation. For instance, upon seeing an animated character fleeing, we have no time to analyze rationally whether there is a soul – we just automatically and inevitably think about its goal in relation to the situation. In other words, the intentional stance taken by the observer is largely an immediate and preconscious response to the liveliness of an animated object.

Interactional and Intentional Liveliness

Hence, the idea of liveliness should at least encompass interactional and intentional aspects. For interactional liveliness the more complex the interplay of forces,

the livelier the behavior is. It follows that a ghost in the arcade game *Pac-Man* is quite lively, because it runs away from Pac-Man who has eaten a power pellet and then turns back to attack him again after the power has gone. The situation also applies to many recurrent chasing sequences in animated films like *The Coyote and The Road Runner*. For intentional liveliness, the movement need not be complex and forces may not be obviously visible. In fact, it is more about surprising the audience by breaking the expected action-reaction pattern. For instance, at the very beginning of the computer animation film *Luxo Jr.* (1986), a plastic ball rolls in and stops next to a desktop lamp, which suddenly “wakes up,” looks closely and pushes back the ball. This act surprises audiences and invokes the intentional stance. Next, the ball bounces back and again, an audience might consider if the ball also has its intention, especially when the ball seems like rolling out of the frame by itself. One does not switch back to the physical stance toward the ball until the lamp looks back into someone in the dark. This first 30-second happening in the film shows us how intentional liveliness meticulously effects based on simple, but unexpected, motion.

Emergent Liveliness

Meanwhile, Arnheim also touched upon other kinds of complex behavior, such as dancing and acting. He sees dance performance as the “dynamics conveyed to the audience visually” (Arnheim, 1974, p. 408). Any movement of body parts must be generated by a “narrow local impulse,” which spreads over all the body and visibly brings about the posture. I believe, to observers, these dynamics are so intricate that the interaction among various impulses can only be perceived collectively. The final image projected to observers is still full of energy, but it is definitely not a simple interactional or intentional behavior. It gives us the third type of liveliness: emergent pattern. This dynamic phenomenon is like the emergent motion of bird or fish flocking, in which each animate being is steering toward local flock-mates and away from overcrowding

simultaneously. Observers cannot tell from where the impulses emerge, because every point of motion is leading, and also being led, now and then, here and there, just like every joint and tendon of a dancer in action. What one can tell is that the overall performance is very lively. Therefore, Dennett's theory may not apply here, because an audience in a dance performance or a beholder of a kaleidoscope seems not to take any of the three stances. One just gazes at the presentation and enjoys the transformation emerging. In fact, this specific kind of liveliness does not concentrate our attentions on the behaviors of particular figures or objects, but rather distribute, or even dilute, our focus over a collective transforming whole. I call this emergent liveliness.

Concentrated versus Diluted Attentions

Emergent liveliness stands quite a different position from interactional and intentional liveliness. Not only does it divert observers from Dennett's three stances, but also it expands our notion of animacy. The animists would argue for this wider horizon, because they attribute life to even falling leaves and waving water. Based on Arnheim's continuum, these natural phenomena are only of low-degree liveliness because they show little or no obvious impulse. Yet, when anime fans are fascinated by the spectacular landscape where grass straws sway in the breeze while the hero just stands still (Figure 3.2), when water ripples spread from the mouse over a button on the computer screen (Figure 3.3), I assert that analysis of only the degrees of liveliness of these phenomena does not truly reveal their vitality. What we need to emphasize here is not the difference in degree, but the difference in kind. There is another kind of liveliness beside those prominent interactional and intentional behaviors. Instead of highlighting tension between forces or surprising acts, this type of liveliness shows a continuous and persistent change in state, which project a life-like ambience to the observers. If prominent liveliness, including the interactional and the intentional, concentrates audiences' attentions on the high-pass momentum, the alternative kind, emergent

liveliness, distributes their attentions to low-pass kinetics. In animation films, the former corresponds to primary motion, like the eye movement of Donald Duck, stretch and squash of a bouncing ball, whereas the latter secondary motion, such as Superman's cape waving in the air and follow-through motion of a ponytail (Figure 3.4). This "secondary" liveliness usually shows rhythmic patterns, such as waving, orbiting, and cycling. Occasionally, it involves stochastic components, like irregularity, chaos, and noise, just as an integral part of real life. In animation practices, these motions are called "secondary" because they are derived from the primary one, but the succession does not mean the order of importance. To the contrary, these ambient patterns are indispensable for a lifelike illusion of life.



Figure 3.2 Genies standing still with their jiggling heads in *Princess Mononoke* (1997)

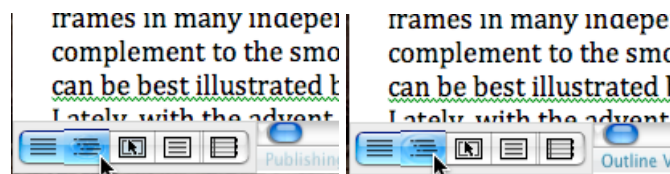


Figure 3.3 Ripples in Microsoft Word

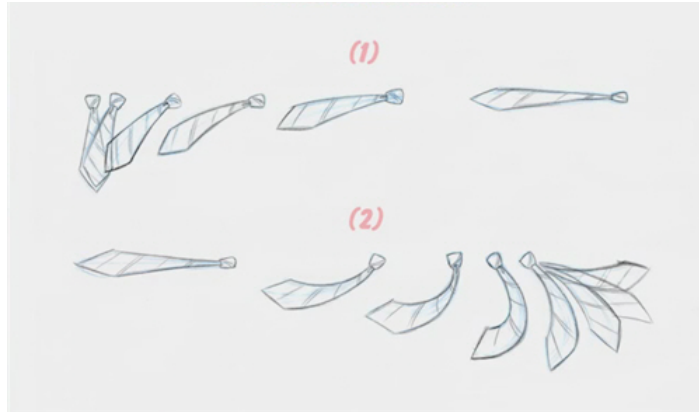


Figure 3.4 Follow-through motion of a tie under the influence of the body movement

The Mechanistic Approach to Motion

The liveliness of movement patterns has been an enduring human concern. Since the Enlightenment, some people have been interested in the means of implementing patterns of movement. Following Cholodenko again (Cholodenko, 1991a, p. 16, 2007, p. 487), here I call them the mechanists. They invented mechanical clocks and various kinds of lifelike automata, such as the mechanical duck apparently able to digest and humanoid robots seemingly able to play flute and to write (Terpak, 2001). All these artifacts exhibit programmed phenomena in patterns. These rhythmic patterns are repeatable because they are outputs of pre-calculated geometric ratios of gears and scales of bars. In short, they are regarded as mechanical motion. Ironically, the primary focus of these automata was to simulate organic behaviors. Similar attempts of highlighting patterns can be found in some non-narrative animation films, like John Whitney's abstract computer animation, in which mathematical algorithms generate visual rhythms, and Oskar Fischinger's *Composition in Blue* (1935), in which movements were programmed according to the soundtrack of the piece. Apart from rhythmic patterns, stochastic phenomena can also be under the spotlight of the show. In fact, chance and randomness have always been greatly appreciated in art practices because they mean less regular, less predictable, and so often more humanlike. This is the reason why the jiggling of frames in many independent

hand-drawn animations is praised as a noted complement to the smooth continuity of commercial animated cartoons. The idea can be best illustrated by the canonical works of Len Lye and Norman McLaren. Lately, with the advent of computer graphics technology, chaos can also be simulated by pseudorandom algorithms. For example, the famous Perlin noise function has been widely used to model many sorts of chaotic patterns in nature, from leaf veins, fire particles, to water waves. The function can even be applied to simulate human-scale jitters in film, such as camera shake, jump cuts, and stop-motion discontinuity.

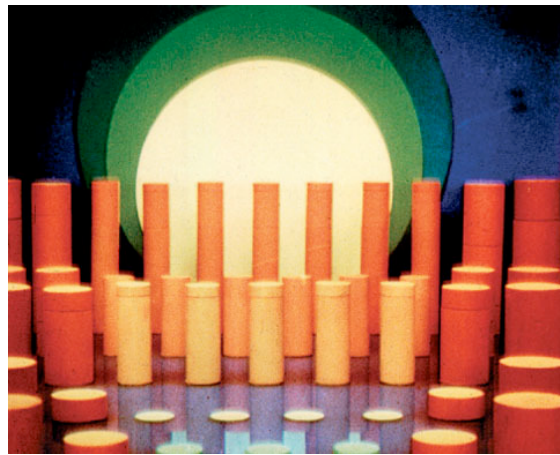


Figure 3.5 A screenshot of *Composition in Blue* (1935)

Summary: Charting Liveliness

In this section, I have described liveliness of three kinds, which result in two different sorts of attentions. Interactional and intentional liveliness refer to those phenomena concentrate observers' attentions, whereas emergent liveliness corresponds to diluted attention.

- *Interactional liveliness* is shown in interplay of force and counter-force. The more complex the action-reaction sequence is, the livelier it is. Observers usually switch between the physical and the design stances.

- *Intentional liveliness* arises from prominent and progressive movement surprising audiences, provoking them into taking the intentional stance. It presents no obvious dynamic tension.
- *Emergent liveliness* can be found in phenomena that dilute our attention across a collective transforming whole. Usually the force pattern is so complex that it becomes rhythmic and stochastic, and it is hard to focus on a particular impulse. The phenomena are emergent in that no apparent consequence can be noticed very often, but they complement the other two types of liveliness to form a complete life-like experience.

In the section “The Observer” of next chapter, I shall show how these various kinds of liveliness are manifested in digital environments, forming a new notion that destabilizes the dichotomies between organism versus machine, and animator versus program.

The Image: Understanding Images

This section examines process of interpreting an image. It starts with describing a trajectory of thought about images, from the literal meaning of the term in the pictorial or visual form, to the metaphorical meaning of the term as a form of mental representation. I then review major structuralist semiotics approaches to images, which tend to separate material images from their meanings, followed by a survey of other ideas that contend that closer connection exists between the material and the mental. Drawing upon insights from cognitive semantics together with recent neuroscience findings both from the field of cognitive science, I argue that animated visual images, which we understand using both our perceptual and motor apparatuses, tie together both material and mental notions of images, constituting an embodied kind of cognitive processes.

“Image” has many meanings. It literally refers to visual images that exist in material forms, like sketches, photographs, or lithographic prints. It also refers to what we see, but do not consider to be fundamentally material, such as shadows or film projections. But the term has also been extended to mean those images that occur in our minds when we think of something, perhaps recalling something seen in the past. Sometimes, the term image can even refer to an imaginative construct shared by a group of people that may be more about emotional impressions and associations, like the general image of a corporation or a rock band. Hence, the term spans both the material and the mental.

Problems in Signification

When we perceive an image, how do we interpret it? Do we “read” an image just as a word or a sentence? How do we make meaning of an image? This kind of question

poses a challenge to many thinkers and researchers in the field of semiotics⁴. Semiotics, also formerly called semiology, is the general science of signs (Barthes, 1973a, p. 9; Gombrich, 2002, Preface), emphasizing the process of signification, which means something “stands for” something else (Eco, 1976, p. 8). The idea developed from concepts from linguistics, in which Ferdinand de Saussure pointed out that for any linguistic unit, there is a link between its concept and its sound pattern. To Saussure, this sound pattern was not physical sound, so not material, but rather a hearer’s sensory impression of the sound (Saussure, 1983, p. 66), or an “acoustic image” as Barthes later called it (Barthes, 1973a, p. 38). Hence, Saussure’s notion of signification refers to a link between conception and perception. Moreover, Saussure asserted that this link is largely arbitrary, which means unmotivated and without any intrinsic natural relation, before being established by the language using community (Saussure, 1983, pp. 67-68). For example, “kot” in Russian and “cat” in English both refer to the same species of quadrupeds – the different sound images are not naturally related to the concept of that species but just “imposed” and “anchored” in the corresponding communities. Once established, these arbitrary links become stable and resistant to change. Although some linguistic signs, like onomatopoeic words, are arguably less arbitrarily motivated at the very beginning, Saussure contended that most of these words have been subjected to substantive phonetic evolution and already lost their originality (Saussure, 1983, p. 69).

Roland Barthes reviewed and extended Saussure’s idea of linguistic signs to apply to many of our everyday objects, such as fashion items and images in popular media (Barthes, 1973b). Yet he distinguished the character of general signs from that of linguistic signs, a distinction that has vital implications for image interpretation.

⁴ Here I consider semiotics to be a field of studies, not a discipline, see (Eco, 1976, pp. 7-8).

First, the image, as a sign, rather than purely a psychological entity, has both material and mental components. In *Elements of Semiology*, Barthes pointed out that the signifier of a general sign has certain material substance because it acts as a “mediator” for the reader to access the meaning/concept of the sign (Barthes, 1973a, p. 47). Since the signifier belongs to the “plane of expression,” the substance of expression can be a material image, for example, a graphic icon, a picture, or a written word. The signified, on the other side of a sign, is a mental image, which corresponds to the “plane of content” and refers to the concept in our minds.

Second, Barthes’s analyses of images problematize the inevitable oppositions of arbitrary and motivated, cultural and natural. For Barthes, an image sign can have both conventional symbolic meanings (connotation) and non-coded perceptual messages (denotation). For example, an image in a spaghetti advertisement is able to show the “Italianicity” of the product, as well as its naturalness and freshness (Barthes, 1977c). In this case, the relationship between image and meaning is both arbitrary and motivated because it has more than one signification, in which the connotation is cultural whereas the denotation is natural. Hence, for Barthes an arbitrary relation is institutionalized, and a motivated one hinges on analogy and similarity between the signifier and the signified.

In fact, Barthes has brought us other questions even more pertinent to the image sign: how can the motivated, and so analogical, nature of images be compatible with the discontinuous signification process? (Barthes, 1973a, p. 52) A photographic image is analogical. It resembles something. The tomato-like object in the advertising image seems like a fresh tomato. This message is not coded, and no one needs any training to recognize it. Barthes calls this kind of phenomena an “awareness of have-been-there” (Barthes, 1977c). This perceptual message can be presented in innumerable pictures of all nuances. How can this type of analogical many-to-one relations be structurally specified by means of signification?

This riddle has been the central question engaging many other thinkers. Nelson Goodman addressed the issue of understanding symbols from a formal approach (Goodman, 1976). He subsumed both visual representation and verbal description under the idea of denotation, or more precisely, classification of samples by means of labels. A label becomes the meaning imposed on a specific group of samples. Why a portrait represents a celebrity, or why a certain style of painting expresses a “blue” feeling, is dependent on the classification system, which is founded on conventions and practices. Ironically, this extremely formal approach brings us back to the culture-nature dichotomy, suggesting that an image is just conventionally linked to its meaning. This view further distances the material image from the mental image by dissolving any intrinsic relationship between the two.

However, not all semioticians agree upon the resolution of the arbitrary-motivated conflict in image signs by the slanted claim of conventionality. Charles Sanders Peirce classified all signs into three categories, namely icon, index, and symbol, which respectively correspond to Firstness, Secondness, and Thirdness in his ontology of triadic relations. To Peirce, intrinsic quality is a first, actual fact a second, habit or law a third. It follows that icons *resemble* something *intrinsically*, indices *relate* to something *causally*, and symbols *represent* something *conventionally* (Hiraga, 2005, pp. 29-31) (see also Stanford Encyclopedia). This triad is generally well received, particularly in visual studies and graphic design, because it seemingly applies well to images. However, there are still ambiguities. For example, a photographic image can be said to be an index rather than an icon as the image is a consequence of chemical reaction to light. Sometimes, it is hard to tell whether a graphic icon of female in a dress is based on resemblance or convention.

Umberto Eco has attempted to develop a general theory of semiotics accounting for all meaning-making phenomena, including the iconic sign. After rigorous and comprehensive analyses, he finally concluded that understanding an image could not be

explained by semiotics alone (Eco, 1976, p. 216). The compromised campaign of semiotics, moving the discipline toward image interpretation rather than a logic or science of images, seems to confirm a central vagueness in the concept of the image – especially when referring to mental images.

Material versus Mental Images: Wittgenstein and Mitchell on Images

Fortunately, there are other thinkers who dismiss the myth of mental images as strictly separate from the material. Ludwig Wittgenstein remarkably pointed out that images inside our brains (mental images) are no more abstract than images outside (material images) because we always think in terms of what we have already perceived, or even what we are pointing at, regardless of whether they are verbal symbols or visual images (Mitchell, 1986, pp. 15,19). To further illustrate the point, W. J. T. Mitchell raised the example of a decoy: when a duck responds to a decoy, it is seeing another duck (Mitchell, 1986, p. 17). To the fooled duck, the material image and the mental image are equivalent (Mitchell, 1986, p. 90). To the hunter, the material image looks “enough” like an imaginative mental image of a duck to fool a duck. In contrast to the verso-recto structure of signs (Barthes, 1973a, p. 30) that has the signifier and the signified on two sides, Mitchell’s argument tends to position material and mental images on the same side with respect to the mind. Both of them are dependent on consciousness. Without a conscious viewer, neither of the two images can exist.

This subjective view of images recentralizes the beholder in the meaning-making process and sheds us new light on image interpretation. An image is no longer an object preceding its viewer, but rather a mental construct based on perception⁵. Images differ from signs in that the structuralist assumption of a pre-existing system plus individual

⁵ This idea echoes Saussure’s perspective of signs, but later structuralist programs have deviated from his original insight.

actualization does not apply. Every image interpretation depends on every perception by every viewer. When we perceive a photographic image, we may “see” a person “having been there.” On seeing a graphic icon on a street sign, we may “construct” a scenario of a car on a slippery road. While we are audience of a puppet theater, we may “imagine” that the puppets grow up to human scale (Gombrich, 2002, p. xvii). These phenomena suggest that understanding images is grounded in perception and cognition. This view on images parallels insights from distributed, embodied, and situated perspectives on cognition, which are in direct opposition to perspectives based in classical Cartesian mind-body dualism (Lakoff & Johnson, 1999). Below, I describe major representatives of these relatively recent trends of thoughts, together with their constructs that will be central to my approach: including material anchors, image schemas, conceptual blends, meshes, and iconic mappings. Each provides an acute contribution to understanding the relationship between images and thought.

Physicality of Images: Hutchins’s Material Anchors

Being one of the major proponents of the idea of distributed cognition, Edwin Hutchins meticulously describes many of our everyday practices and instruments, such as queuing and analog timepieces, in which physical structures of images represent elements for conceptual processes. He asserts that material structures and patterns, like marks or diagrams, can provide us with stable images for complex mental operations, such as calculation done on paper or navigation with a compass. He calls these images material anchors, which “hold” information in place or incorporate constraints for mental manipulation (Hutchins, 2005). It follows that a material image, for example, marks on paper or a compass rose and needle, can act as a direct input to the cognitive process, in which the world is used as its own representation. In Wittgenstein’s words, we compute by manipulating the marks or diagrams in our minds. While Hutchins’s analyses cover mostly goal-specific, instrument-oriented tasks, the material image of each artifact or

instrument is equivalent to the mental image for manipulation. The major advantage of anchors to cognitive processes is that one can reduce memory and processing loads by building the constraints of the task into the physical structure of the artifact.

Spatiality of Images: Image Schemas

While material anchors represent the distribution of concepts onto physical structures of images, image schemas reciprocally suggest that many concepts are built upon our experiences with spatial structures of perceptual images. Based on numerous examples in our everyday use of language, George Lakoff and Mark Johnson shows that metaphors not only allow us to communicate abstract concepts by projection of similarities but actually structure largely our ways of thinking through entrenchment (Lakoff & Johnson, 2003, p. 6). Many of these metaphors are based on our bodily and perceptual experiences in space. For instance, when we say “time flows,” it metaphorically seems that time is a fluid substance flowing by us and we are observers (Lakoff & Johnson, 1999, p. 158). These metaphors are so conventional and entrenched that they often just gone unnoticed, and the corresponding mental images are so embedded in our minds that they exist as extremely skeletal and schematic images, what cognitive semantics calls “image schemas” (Lakoff & Turner, 1989, p. 97). Lakoff and Mark Turner differentiated image schemas from image metaphors in that the former are not mental images but instead very general spatial structures (Lakoff & Turner, 1989, p. 99). In other words, only structural patterns are preserved in the schemas for spatial reasoning.

Integration of Images with Concepts: Conceptual Blends

Gilles Fauconnier and Turner have elaborated the insights of conceptual metaphors and mental spaces (Fauconnier, 1985) to introduce conceptual blending theory (Fauconnier & Turner, 2002). The theory describes a basic mental operation that

generates new meaning by integration of concepts. The operation constructs a partial match between multiple input conceptual spaces and selectively projects from those input spaces into a novel “blended” conceptual space. Conceptual metaphors are a kind of asymmetric blends in which the “source” input provides the topology for the understanding of the “target” input that is completely project into the blended space (Fauconnier & Turner, 2002, p. 127). Blending is a dynamic process, and successive blends give rise to an emergent integration network, which is pervasive in everyday life, as well as other creative feats like rhetoric, reasoning, gameplay, and even interface design (Fauconnier, 2001, p. 279). Fauconnier and Turner’s analyses also are sympathetic with Hutchins’s distributed cognition analyses as they examine everyday objects like watches and money to illustrate how our minds interact with the world. Their discussion even extends to cover written, spoken, and sign languages (Fauconnier & Turner, 2002, pp. 210-216). To this, Hutchins’s response is not without doubt. He contends that the arbitrary relations in most linguistic signs make them a very weak type of material anchor, because not much analogical information is held in the material form of these signs. Each of them just provides a distinctive perceptual identity from one another (Hutchins, 2005). To Fauconnier and Turner, a material anchor can just be a structural constant for a concept, like money notes. In contrast, Hutchins demands more information loaded onto material anchors. He thinks the compass rose, with all of its marks, is more important than just the needle, which he believes Fauconnier and Turner did not emphasize (Hutchins, 2005).

Matching of Images with Bodily Experience: Glenberg’s Meshes

Although there are nuances between Fauconnier and Turner’s interpretation and Hutchins’s view, material anchors for conceptual blends mark an indispensable link by which part of human memories can be projected onto world objects. In fact, Arthur M. Glenberg has also investigated the connection between memory and the world based on

Lakoff and Johnson's thesis of embodied mind (Glenberg, 1997). His article repositions memory as cognitive apparatus to combine, or in his words "mesh," perceptual patterns projected from the environment with patterns of interaction from bodily experiences. The two patterns are compatible because they are both "encoded" in terms of one's body. One can recognize a walking path as the "path home" using a match of patterns between the perceived environment and embodied motor knowledge in one's body memory. If material anchors suggest a "download" of structural information from memory to artifacts by perceptual processes, Glenberg's notion of mesh recalls an "upload" of spatial and functional meaning from the environment to memory through embodied interaction. The two ideas highlight different portions of a mind-matter continuum, but they definitely do not draw any boundary. Instead, they mobilize interplay between mind and matter through the body.

Mapping between Images and Metaphors: Hiraga's Iconicity

For image schemas, material anchors, and meshes, images, no matter material, perceptual, or mental, all play an important role in providing topological or structural patterns. However, there are still other qualities of images that cannot be overlooked. Masako K. Hiraga's approach to iconicity in poetry considers a wider spectrum of iconic representation based on Peirce's well-known icon-index-symbol trichotomy of signs. Hiraga particularly looks into Peirce's further division of icons into images, diagrams, and metaphors, which are delineated with increasing abstractness from concrete resemblance to structural and relational analogy (Hiraga, 2005, pp. 30-34). While diagrams and metaphors are mostly addressed by other aforementioned theories, Hiraga's analyses also cover the most immediate and intrinsic relations, such as the shape of a poetic text resembling a swan (Hiraga, 2005, p. 101).

To Hiraga, iconicity is the mapping between form and meaning, in which meaning can be a mental image whereas form can refer to either imagic or diagrammatic

quality of a text. Hence, an imagic mapping relates objects by immediate resemblance and a diagrammatic mapping by structural analogy. These mappings, like metaphorical projection, are subsumed in conceptual integration networks to explain how a reader generates multiple mental images, or in Hiraga's terms "iconic moments," from a text. The value of Hiraga's iconicity to image interpretation is bringing the most intrinsic qualities of images back to the nexus between perception and cognition.

Summary: Image Interpretation

After a survey of all these major thoughts from a vantage point of image interpretation, I arrive at the following points:

- *Immediate-Arbitrary Relations of Images*: Images differ from verbal signs in that intrinsic, immediate, and analogical relations co-exist with arbitrary and conventional relations.
- *Relating Material and Mental Images*: Material images and mental images should be seen as in relation rather than in opposition. Both of them should not be analyzed without the beholder because they both are mental constructs. In some cases, the two can even be equivalent.
- *Connecting Image Perception, Cognition, and Bodily Interaction*: Images, along with their material structures and perceptual patterns, can take part in cognitive processes like matching. Moreover, the mind is capable of matching patterns projected from perceptual images or resulted from bodily experiences.
- *Mapping Form and Meaning of Images*: The mind can recognize various degree of iconicity of images, which maps form and meaning along a continuum from intrinsic to abstract qualities.

However, none of the above theories explicitly address animated images. When images are moving, would we understand or even "sense" them in a more immediate way? Would animated visual images be closer to mental images? How would animated

images be matched with other cognitive patterns in conceptual networks? How would animated images be cognitively connected to bodily interaction?

Animated Images as Embodied Images

Hutchins's arguments for material anchors mainly focus on human-performed instrumental and operational tasks, so his material images have to be stable and faithful representations of the elements to be manipulated in the cognitive process. However, this faithfulness does not necessarily apply to cases in which the outcome is not a priori clear and task specific, such as process-driven imagination-laden creative activities. For example in filmmaking, a storyboard is not strictly a faithful representation of the mental image inside the director's mind, but is also a device used for contingent reflection on a creative work-in-progress, projecting evocative sensation onto the work that goes beyond the physically represented information, and allowing that work to trigger generation of subsequent imaginative images. W. J. T. Mitchell, based on other influential thinkers, has pointed out that words can be so vivid that provoking imaginative images, in his words "verbal image," even livelier than the original sight (Mitchell, 1986, pp. 23-24). I believe that animated visual images transgress the boundary between the original and the imaginative even more vigorously, mobilizing viewers' motor-sensory connections and constituting embodied understandings of sensations.

Consider that the storyboard of a film in progress projects the director's approximation of the intended outcome. At some point, the director will need an animated visual image, technically called a "rehearsal," an "animatic," or a "rough cut," especially when one wants to elicit visceral sensations such as disgust, sorrow, nervousness, and others. There are many nuances to these "gut feelings" that static images may not be able to convey. Instead, they have to be performed as actions in animated images. For example, a viewer is able to distinguish an animated character's giggling from trembling, because the viewer perceives and understands it as exhibiting

lifelike qualities previously experienced in her or his everyday life. This is quite different than the use of culturally specific symbolic conventions (such as trembling lines) without which a still image could not convey these distinctions.

While being careful not to overstate these claims, results regarding the activation of brain structures known as mirror neurons have been posited as suggesting that, when perceiving a performed action in a moving image, the viewer's interpretation relies upon evocation of the corresponding sensorimotor knowledge from a repertoire of her or his own embodied experiences (Rizzolatti & Sinigaglia, 2008, pp. 94-98, 124-125). The coupling of perception and action enables the viewer to "recall" the associated sensation. When we see someone jumping restlessly, we understand one's excitement, not by "reading," but rather "sensing." In short, animated images constitute an embodied cognitive process. By the same token, a movie director can "feel" whether an actor's performance is matching his mental image, or an animator can "detect" if her or his character is moving right. This kind of visceral understanding largely takes place at the immediate, or even preconscious, level – requiring minimal cognitive effort.

Cognitive semantics research also provides us with accounts of understanding sensation through perception. When discussing animacy, Mark Turner states that we cannot perceive others' sensations, so we can only infer their sensations by comparing their actions to our own reactions in similar situations (Turner, 1996, p. 21). He refers this analogical inference to a type of parabolic projection, or metaphor, in which partial structure of a source story of the perceiver, including action and emotion in particular, is projected to a target story of the perceived.

It follows that the director or animator can infer the sensation by cognitively projecting her or his own experience to the perceived action. This act of "inference" seems to be suggesting that the projection takes place at a higher cognitive level, demanding conscious mental operation. In fact, some mental projections can be cognitively effortless. This point can be illustrated in the terms of conceptual blending

theory. The matches between two input mental spaces, what Fauconnier and Turner call “vital relations” (Fauconnier & Turner, 2002, p. 92), of some blends can be so tightly compressed that they become automatic and unnoticed. Fauconnier has cited the computer interface phenomenon as an example of this kind of immediate blends (Fauconnier, 2001, pp. 264-265). In this regard, I would like to add that, for the director or animator, there is also a blend of a past experience and a perceived action yielding an inferred sensation. The compression can be so tight that the animated image is immediately associated with the sensation. The image becomes an immediate and embodied input to the integration network.

As mentioned earlier, Hutchins coined the term “material anchor” to mean those material objects or images with stable patterns and structures “locking down” specific information in input concepts. I believe that animated visual images, as embodied images, not only “hold” information but also “embody” sensation or meaning, which trigger imaginative elaboration in blending. Instead of concentrating on goal-specific computation, I shall explore material-based open-ended imagination through cases of fluid and flexible representations in the form of animated visual images in the section “The Image” of next chapter.

The Body: Embodied Meaning of Action

In this section, I look into how observers engage with animated artifacts. Regarding the traditional film form of animation, audiences are mainly engaged through perception, except for some subtle body movements like shifting between reclining or upright positions. In digital media, the engagement involves not just sensory perception, but also motor-based bodily interaction. An observer absorbed in audio-visual content, for example in a multimedia website, is also driven to take timely action or to react promptly, because the animated artifact invites participation meanwhile continuing its transformation. The body of the observer is engaged continuously, in terms of attention – like being captivated in a theater, and also alternately, in terms of action – like participating in a tennis game. Most important of all, the observer makes meaning through this engagement of sensory and motor apparatuses with the system, echoing the central embodied cognition perspective that meaning-making processes involve both perception and motor action. On the sensory side, animated visual images constitute embodied understanding of sensation, as I have argued in the previous section. On the motor side, as I shall discuss in this section, bodily motor action embodies intention. The two sides conjoin a motor-sensory feedback loop between the system and the user. I believe animated systems entailing this motor-sensory connection are able to give users an embodied experience of an expanded illusion of life.

To illustrate this new notion bridging animation and motor action in animated systems, I first review major underpinnings related to motor-sensory feedback in physical or machine-mediated environments, followed by insights from phenomenological thoughts of how users make meaning of bodily interaction with these environments. Then, drawing upon Maurice Merleau-Ponty's perspective on bodily action, which suggests that our bodies are able to turn movement into meaning in the form of motor habits, I propose considering motion-based user input, in which meaning is embodied in

the qualities of motion, as a preferred alternative to the conventional point-and-click method of user input for animated systems. In such systems, as multimedia websites or entertainment applications, the motion components persisting in both motor input and animated feedback enable a tighter motor-sensory connection and give rise to more embodied and evocative meaning reminiscent of our everyday motion.

Meaning in Environments: Feedback and Affordance

Various motor-sensory circuits persist in interaction between human users and engineered systems in different scenarios. When riding a bike, typing on a keyboard, playing computer games with a gamepad, and using other motion-sensing input devices, a user has to engage part or all of the body in “controlling” the machine and “sensing” the feedback. The interaction mechanisms are not unlike how an animate interacts with its physical environment, for example, a bird subtly adjusting its feathers to round about approaching obstacles in the middle of its flight. Quite a few scholars have paid attention to the similarities in interaction mechanisms between engineered systems and physical environments. Norbert Wiener is credited with introducing the term “cybernetics,” which from the Greek means “steersman,” to call the study of control and communication theories in animals and machines (Wiener, 1961, p. 19). To Wiener, the steering engine of ships is one of the earliest and best-developed forms of feedback mechanism. Although his interpretation of “feedback,” like steering, is inclined to regulating output by adjusting input (Wiener, 1961, p. 13), which is slightly different from my emphases on lively reaction from systems, his awareness of the value of animal reaction to nature reminds us that observations of the natural environment and its inhabitants is informative to designing interactive systems because of the forgotten commonalities between two sets of environments. Recall the first time we attentively used the mouse to drag a file to the trash bin in a GUI environment. We kept moving the mouse while focusing on the moving file icon until it reached the trash bin. This sensorimotor phenomenon is

evocative of our physical experiences of throwing away items in the office. We need to “see” something being “put” in the bin. In the GUI environment, after being so attentive at the first trial, we have gone unnoticed of the details in repeating the action. Today we are fairly used to operating a wide array of input devices and interpreting sensory feedback from various systems, just like how our bodies get used to the physical environment. Hence, digitally mediated environments absorb the body no less than the physical environment.

In addition, J. J. Gibson’s theory of affordances also marks the vital link between bodies and environments. Gibson coined the term “affordance” to mean those environmental features that support what animals can do. Gibson believes that these properties in an environment, whether it is the natural or the artificial (in fact Gibson dismisses this dichotomy in environments, see (Gibson, 1986, p. 130)), are usually directly perceivable, without requiring much learning (Gibson, 1986, p. 143). This idea inspired Donald Norman’s asserting that in designing interfaces affordances *should be* made perceivable to users (Norman, 1988, pp. 22-23). Norman’s proposal then has given rise to the prevalent user-centered design paradigm, which aims to make interfaces easy to use and to learn. In his well-known book *The Psychology of Everyday Things*, Norman emphasizes visible, natural, and perceivable clues laid out in the environments. To make interfaces user-friendly, one has to make those clues easy to notice. How? Here Gibson again sheds us some light. As Gibson puts it, “affordances are properties taken with reference to the observer” (Gibson, 1986, p. 143). A property that is “commensurate with the body of the observer” is “more easily picked up.” Therefore, Gibson’s idea also reminds us that the relationship between bodies and environments should be recentralized in interaction and interface design.

Embodied Meaning in Tools: Dourish’s Embodied Interaction

With an awareness of the role of the body in designing interactive systems, Paul Dourish advocates an approach to interaction design grounded in the idea of embodiment. Dourish draws upon notions from phenomenology, particularly Martin Heidegger's "Being-in-the-world," to interpret embodiment as people's engagement in the world in order to make meaning of it (Dourish, 2001, p. 126). In other words, meaning is embodied in engagement. For Dourish, engagement includes both physical and social interactions. Therefore, to create embodied interactive systems he suggests making use of our "familiarity" with the mundane everyday world, including practical experiences with physical objects and communication skills in social communities. His book delineates several high-level principles to achieve this goal.

By embodiment Dourish seems to mean largely the embodiment of tools in the working environment. This bias toward production may have originated from Heidegger. As in his well-known hammer analogy describing that a carpenter would only see a hammer available for his tasks (what he calls "ready-to-hand"), Heidegger presupposed a "work-world" in his pronounced dictum (Moran, 2000, p. 233). It follows that Dourish's proposal also emphasizes the practical and functional meaning of everyday world objects and tools. Although this meaning of "everyday experience" with objects is in line with the major notion of phenomenological embodiment (Lakoff & Johnson, 1999, p. 36), Dourish tends to concentrate on objects and practices, but sometimes overlooks the body that uses those objects. His approach focuses on the ways that objects, artifacts, or equipment, are put into use, but seldom elaborates on how our body acts upon and reacts in these situations. In fact, Heidegger's "Being" is more about things like equipment rather than human beings (Moran, 2000, p. 412).

Embodied Meaning in Body: Merleau-Ponty on Bodily Action

Heidegger believes that meaning is embodied in practice and in action (Dourish, 2001, p. 184), but he largely ignored the role of the body in the process. In contrast,

Maurice Merleau-Ponty thinks that “our body is our general medium for having a world” through habit (Merleau-Ponty, 1962, p. 146). By “habit,” Merleau-Ponty means specifically motor habit. Based on Gelb and Goldstein’s empirical studies of brain-damaged patients, Merleau-Ponty describes two types of bodily movement, namely “Greifen,” meaning “to seize,” and “Zeigen,” as “to show” (Merleau-Ponty, 1962, p. 123). The former refers to the kind of concrete movements toward some existing objects, while the latter the kind of abstract movements not relevant to any actual situation and dependent upon one’s imagination. For example, an impaired patient may be able to touch his nose because of a mosquito sting, but not able to point to it without preparatory movement. A typical subject can perform a salute, but the patient cannot without placing himself in an actual situation. It seems that the patient cannot enter into an imaginary situation. In fact, unimpaired people have turned most abstract movements, like pointing or other gestures, into motor habits. Merleau-Ponty believes that it is our body that actually absorbs meaning, in the form of bodily experience (Merleau-Ponty, 1962, pp. 146-147). In Dreyfus’s words, our body gets involved through practices and acquires bodily skills (Dreyfus, Wrathall, & ebrary Inc., 2006).

General Motor Input: Direct Manipulation versus Gestural Input

The two types of bodily action mentioned by Merleau-Ponty aid in articulating a general divide between mechanisms of motor input in HCI, where the two sides are direct manipulation and gestural input. The former corresponds to “grasping” whereas the latter “showing.” (While this separation is at the mechanism level, another contention exists at the design level, which is between direct manipulation and intelligent agent, and is covered in Chapter 7.)

Ben Shneiderman is credited with summarizing the central ideas of direct manipulation, which include spatial representation of objects, physical action, and visible feedback. In other words, they are common phenomena prevalent in the graphical user

interface of today. As Shneiderman puts it, the primary goal is to bring human-computer activity back to the level of early stages of Jean Piaget's theory of child development, in which children comprehend largely through physical actions. Interfaces based on this principle would be easier for children and adults to use (Shneiderman, 2003 [1983]). That is to say, direct manipulation requires not much learning and just demands users of the innate capabilities of concrete movements.

However, the use of gestures in HCI supposes that users perform abstract movements with their empty hand such as pointing, alongside other modes of input like speech. One of the first experimental interfaces incorporating these is *Put That There* (Bolt, 1980), in which a user is able to move objects on the graphical interface by a combination of pointing and ordering. The act of pointing in this work differs from using conventional pointing devices like the mouse. The former is not an example of direct manipulation, because it does not involve the act of dragging (Bolt, 1980). Instead, general pointing is a gesture, although a very primitive one that has gone unnoticed. Alan Wexelblat quotes Adam Kendon, a leading expert in gesture study, that a gesture must have features, motion or shape, that make it identifiable (Wexelblat, 1994). It follows that a click-and-drag action with a mouse is not a gesture, neither pressing a key on a keyboard, because the corresponding hand movement and shape have nothing related to the meaning and consequence of the action. The speed or path of dragging the mouse or pressing the key has no significant effect on the result. In fact, Wexelblat and other researchers like Francis Quek are more interested in empty-handed natural gestures and how these can augment multimodal input. That means they look for interfaces other than conventional input devices. However, I find that their rigorous reviews of gesture taxonomy are still informative for analysis of gestural input enabled by emerging input technologies including touch-based input and motion-sensing hand-held devices.

From Gesture-Based to Motion-Based Input

Wexelblat surveys and summarizes gesture types in an array of symbolic (by convention, e.g., thumbs-up-for-good), pantomimic (miming the use of objects and making similar motions), iconic (hand-acting), deictic (pointing or directional), and rhythmic (Wexelblat, 1994). In the context of human-computer interaction, to select by conventional pointing devices like the mouse can be seen as a mediated type of deictic gestures. In addition, many touch-based input devices like the touchpad on many laptop computers facilitate a more natural type of pointing without the intermediate. The touchpad and touch screens open up possibilities of applying more gestures as motor input, such as to select a group of objects by circling on the touchpad, to turn a slide by swiping, or to rotate a picture by a pinch-and-turn action. All these belong to the overlapping area between symbolic and pantomimic gestures. Meanwhile, motion-sensing input devices enabled by accelerometers or other motion-tracking technologies also make possible input between pantomimic and iconic gestures. For instance, a user can grip the controller of a video game console just like swinging a tennis racket, or he can shake the cell phone to undo a step and tilt it to steer a vehicle in a racing computer game. It seems that pantomimic gestures are particularly promising in these emerging modes of input, which are more natural, situated, and embodied to users because pantomimes always rely on pose and more importantly motion to show the use of some “phantom” objects, that is, embodied meaning. According to Gordon Kurtenbach and Eric A. Hulteen, the motion of a gesture must be significant to its meaning (Kurtenbach & Hulteen, 1990, p. 310). Hence, bodily action and motion are crucial to convey meaning in gesture-based interaction.

Embodied Meaning in Motion-Based Input

Merleau-Ponty posits that what is damaged in the patient is the “power of laying out a past in order to move toward(s) a future.” He calls this the “intentional arc,” which is our consciousness toward something (Merleau-Ponty, 1962, pp. 135-136). The

intentional arc works beneath the levels of conscious conceptualization and perception, through which we acquire bodily skills and build motor habits. Our bodies “absorb” motor knowledge and take care of our everyday motion (Dreyfus, et al., 2006). Therefore, a typical subject can perform seemingly imaginary abstract movements as easily as concrete movements without a thought. Merleau-Ponty believes that motility reveals our consciousness as “not a matter of ‘I think that’ but of ‘I can’” move toward something (Merleau-Ponty, 1962, p. 137). In other words, bodily motion represents intention and meaning. This idea applies to not only gestures, but also motor action at large. Hence, I argue against any clear distinction between the two general types of bodily movement, “to seize” and “to show”, and so the two types of motor input in human-computer interaction, for practices would enrich our intentional arc (Dreyfus, et al., 2006), turning any possible movements to motor habits. For example, double-clicking a mouse button is a gesture at first, but repeated exercising turns it into a habitual action embodying a user’s intention to confirm. Hence, stringently separating gestural input from direct manipulation might not lead us toward tighter motor-sensory connection and more embodied meaning in animated systems. Instead, distinctions should be drawn based on whether a motor input contains meaningful components to our body, that is, situated, embodied, and evocative motion. Consider the jog dial of a video tape recorder (VTR). When a user spins the dial, the motion, including speed and direction, convey his various intentions of going forward or backward at variable speed. This kind of motion has already been borrowed as a gesture of hurry in live performance, in which the spinning speed roughly reflects the hurriedness! (See Figure 3.6) Similar motion-based signals can be applied to computer interaction as well, such as on the touchpad. Therefore, I call for a specific kind of motor input, which is motion-based input, in which motion embodies the intentional arc with meaning and information. In short, bodily motion is a spatiotemporal embodiment of meaning, because it manifests the intentional arc enveloping past and future, here and there. When one performs a bodily action, the body part, such as hands

or fingers, moves in space from time to time, carrying and conveying one's intention. I believe interactive systems taking motion components into consideration result in tighter motor-sensory links and entail more embodied meanings.



Figure 3.6 Use of the jog dial of a VTR

Summary: Embodied Meaning of Bodily Interaction

Based on the above review of motor action in the context of digital interactive media, I come up with the following points about the embodied meaning of bodily interaction:

- *Body Adapting the Digital:* Based on Wiener's notion of feedback and Gibson's concept of affordance, I deduce that bodily interaction of an animate being in the physical environment draws a parallel to that in digitally mediated environments, and our body gets used to digital environments just like we adapt to the physical world.
- *Action Embodying Meaning:* Reviewing Dourish's embodied interaction approach, I argue that meaning is embodied not just in the ways we use objects but also in how our body acts upon artifacts.

- *Motion Revealing Intention*: According to Merleau-Ponty's philosophy, our body can turn any movements into motor habits, which reveal our intentions toward something.
- *Spatiotemporal Embodiment*: Meaning is best embodied in bodily movements taking place in both space and time.

Therefore, in order for users to make embodied meaning in interacting with digitally animated artifacts, interfaces should incorporate more motion-based motor input. On the one hand, this kind of input, as I shall discuss in the section “The Image” of next chapter, couples with animated feedback in a motor-sensory loop embedded in digital interactive media, usually resulting in adaptive, evocative, and imaginative meaning. On the other hand, in the section “The Body” of the next chapter, I shall elaborate on the idea of spatiotemporal embodiment, and show that motion-based input, together with interactive generative animation, would lead us to digital environments entailing such kind of embodiment.

The Performer: Live Performance of Animation

In digital environments, generative and interactive animation, rather than like film-based movies, have more in common with puppet performance or automaton demonstration. Instead of just viewing direct playback of recorded moving images on the screen, audiences in front of a computer screen or a performance stage both see presentation images that can show possibly different versions on every viewing, as well as immediate variations in response to their occasional actions. This section delineates the ontological and phenomenological similarities between puppet performance and the new animation paradigm that I have been arguing for. It starts with a brief comparison of animation and puppetry by way of the overlap in physical shadow puppets. I then describe how computing technology enables performance animation in virtual space. Finally, I show that with advances in real-time control and rendering technologies animation phenomena in the digital age are now rendering live performance a new meaning.

Animation versus Puppetry

Animation and puppetry are commonly viewed as coming from different lineages of arts. The former is a subcategory of film, a visual art. The latter is a marginal type of theater, a performing art. Puppetry has very historical and even legendary origin in several ancient cultures, such as China and Java. In contrast, the birth of animation is always attributed to those nineteenth-century European bourgeois entertainment toys like the zoetrope and the phenakistoscope. Apart from their contrasting origins, the two arts also have obvious differences in terms of media materials, technologies, and practices. One is traditionally based on animators' frame-by-frame manipulation of drawings, pictures, or images, while the other relies on puppeteers' direct manipulation of tangible objects or doll-like figures.

However, there is ambiguity in certain niche areas. The noted silhouette filmmaker Lotte Reiniger has a book called *Shadow Theatres and Shadow Films*, juxtaposing the two arts with a focus on her life-long favorite medium: the shadow puppet (Reiniger, 1970). A glimpse of the illustrations in the book reveals a remarkable connection between puppetry and animation. The articulated figures used in two domains are basically the same. Since shadow-figures are usually flat, an animator can take one from the shadow theater without the rods, lay it flat on a glass plate, and shoot it with an overhead camera, just like an ordinary filming setting for cel animation (Reiniger, 1970, p. 85). As Reiniger calls animation as “an entirely new kind of puppeteering” (Reiniger, 1970, p. 82), to the puppeteer the two arts have ontological similarities. Moreover, the projection of silhouette films, for example, the Reiniger’s classic *The Adventures of Prince Achmed* (1926), in the cinema looks very much like a performance in the shadow theater (although the former is rear projection while the latter front projection). From the audience’s viewpoint, the two media also have phenomenological intersection where both give life to some tangible inanimate objects. Hence, the ontological and phenomenological likeness brings us an ambiguous term, puppet animation, which can refer to films created with stop-motion filming techniques on tangible figures, or those directly recorded from live performance of tangible figures.

Performance Animation: Tillis’s Double Meanings of “Real Time”

The advent of so-called “computer animation” seems to further blur the distinction between animation and puppetry. With latest advances in various real-time control and rendering technologies, computer-generated characters become comparable to puppets, and the term digital puppet is used not just as a simple metaphor. For instance, by means of motion capture, a computer animator can enact a computer-generated version of Mickey Mouse in virtual space just like a puppeteer wearing a physical all-encompassing bodysuit (performers in all-encompassing bodysuits are

theoretically regarded as puppeteers, see (Kaplin, 2001)). In his attempt to reposition puppetry in the digital age, the performance studies scholar Steve Tillis has explicitly drawn the correspondence between motion-capture techniques and all-encompassing bodysuits, as well as the analogy between inverse kinematics in computer graphics, which allows an animator to use a handle to move the end of an articulated chain so that each joint will rotate accordingly, and the strings or rods in puppetry (Tillis, 2001). Since data from inverse kinematics or motion capture can be processed and rendered as computer graphics characters in real time, today's animators can see the resultant movement according to their control instantaneously, making the operation of a digital puppet strikingly similar to that of a physical puppet. This synchronicity of operation makes computer characters resemble their physical counterparts in the way that stop-motion puppets are unable to match. Although digital puppets and physical puppets inhabit different spaces, computer rendering on the screen provides a visual experience comparable to what we see in the theater. Hence real-time control and rendering technologies fulfill what Tillis considers the double meanings of "real time" (Tillis, 2001). On the puppeteer or animator side, it refers to a synchronicity between control and resultant movement. On the audience side, it in addition includes real-time reception as well. Both sides together suggest animation and puppetry can be put in the same category.

Other intriguing examples showing digital puppets include the online application kit *Sodaconstructor* (www.sodaplay.co.uk) for constructing animated models and the program *Animata* (animata.kibu.hu) designed to create interactive animations for the theater. *Sodaconstructor* allows users to make digital creatures composed of point masses and articulated by elastic skeletons. A user can drag a point mass and the whole creature body will move along, like a puppeteer moving a rod attached to a shadow puppet. Meanwhile, *Animata* supports the performance of a reverse shadow play, in which a player standing in front of a clear background uses the whole body to control a computer-

generated Indonesian shadow puppet projected facing him. The puppet is like a full-sized avatar of the player. Both cases demonstrate a kind of “performance animation” enabled by digital interactive technology.

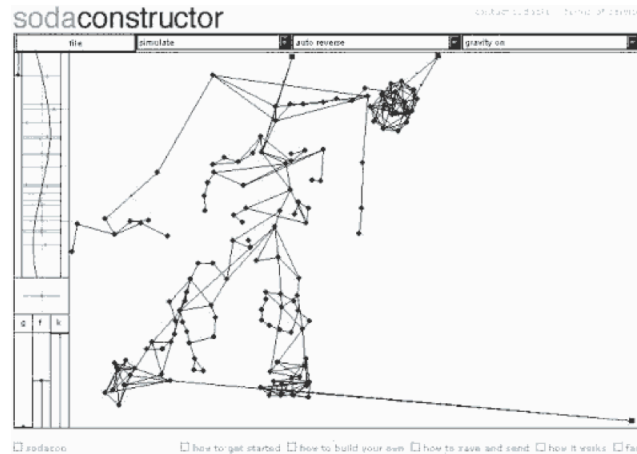


Figure 3.7 A screenshot of Sodaconstructor



Figure 3.8 A digital shadow theater using Animata

For Tillis, the term “performance animation” applies to any computer-generated figures operated by animators and received by audiences in real time (Tillis, 2001). While Tillis primarily focus on anthropomorphic figures like puppets, I add that the term should apply to any computer animation as long as there is a synchronicity of operation and reception. For instance, the Apple iPhone entertainment application *Spawn* generates real-time particles moving like swarms, meanwhile a user can lead the movement with a

finger moving on the touch screen, trigger firework explosion by a tap, and others. Although the digital “puppet” here is non-representational and totally abstract in form, it does “perform” according to the user’s finger action and is received by the user’s eyes simultaneously.



Figure 3.9 Snapshots of the iPhone application Spawn

Observer and Performer: A Mutual Sense of Liveliness

In above cases, computer systems invite user participation by means of real-time control and rendering technologies. Users are responsible for part of the happenings just like performers. Yet, some might argue that participating users are audiences and are by no means performers. Although a performer and an audience might see a similar visual content (e.g., in the shadow puppet show, the audience views in front of the screen while the puppeteer watches behind it, and what they see are mirror images), what they really perceive may be drastically different. It is because they have different selective attention. As I have shown in the section “The Observer” about perception of animacy, the observer perceives a holistic illusion of life, including prominent behaviors and emergent patterns. In contrast, the performer has to pay most of the attention on the nuances in control and feedback, which result in a slightly different composition in the liveliness perceived. The

performer usually more attends to patterns than behaviors. In his famous essay about the art of marionettes (Kleist, 1810), Heinrich von Kleist describes a fictive dancer who admires the unconscious and natural movement of puppets, which always follows gravity faithfully. Kleist thinks that humans' self-consciousness is upsetting the "natural grace" of movement. To the performer, the liveliest movement hinges on not just internal impulse but also natural resonance, that is, body rhythm and coordination. Paul Wells also makes a similar argument, for Wells many works of abstract animation are very animistic because the pure movement depicted is a reflection of the automatic and preconscious coordination of the eye and the hand of the animator (Wells, 1998, p. 32). In fact, this preconscious motor-sensory loop is the ground of many animation productions. Generally, both performers and animators favor what I call diluted attention. Therefore, animators and audiences basically take different perspectives on liveliness.

However, today's real-time technologies bring about a mutual sense of liveliness between the performer and the observer, making animators and audiences alike in digital media. For example, many authoring environments provide animators with real-time animation previews, which inadvertently turn an animator into the audience of her or his own work in seconds, like a puppeteer in shadow play. Indeed, instantaneous reactive features of many multimedia systems allow viewers to feel the preconscious loop of motor action and sensory feedback in a similar way as what animators have always experienced. As soon as one is engaged in the transient and continuing interaction with the system, one has to make timely response, and attention will be fully occupied by the motor-sensory coordination. The participants, including the animator and the audience, of a work of interactive generative animation, are becoming both the performer and the observer. They would understand and share each other's perspectives of liveliness through the real-time interactive system, which blends their roles into a new form of live performance.

Puppeteerly Animation

Real-time systems transcend the new animation phenomena as performance animation. This is a new notion blending the traditional concepts of puppetry and animation. I call it “puppeteerly animation,” as opposed to the traditional film-based “viewerly animation.” This neologism follows similarly the rhetoric of Roland Barthes’s “writerly text,” in which the reader produces, instead of consumes, meaning of the text like the writer. By the same token, in puppeteerly animation, the viewer creates, instead of receives, animation phenomena like the puppeteer in a live performance. Yet, some people might think that this kind of performance animation should not be regarded as live, because the leading creator is always absent from the show. For them, a theatrical puppet show is live but a videotaped puppet show or a stop-motion puppet animation is not, because the puppeteer is absent during audience reception. Within this view, it follows that puppeteerly animation cannot be live since system designers, programmers, and animators are all absent during user interaction. I argue, in contrast, that automaton demonstration, like the mechanical duck or flute player invented in eighteenth-century Europe, or the tea-serving *karakuri* created during the Edo period of Japan, is contrastingly regarded as live performance, although the designers or inventors are also commonly absent from the show. It shows that spatiotemporal co-presence of the creator and the observer is part of the ritual in the traditional stream of the performing art only. It is not a necessary condition of liveness.

Regarding the issue of presence in performance within today’s mediatized context, Steve Dixon asserts that the essence of presence is no longer about physical presence, but rather audience interest and attention (Dixon, 2007, p. 132). This view is best demonstrated in digital media, in which time fragmentation (e.g., video-on-demand service on interactive television or web-based television, online weblogs, instant messaging service, etc.) and telepresence (e.g., online chat-room service, tele-videoconferencing, social networking websites, virtual worlds like SecondLife, etc.) are

commonplace. Immediacy and proximity are no longer defining characteristics of liveness in the digital age. In other words, computer-mediated performance animation engaging its audiences, both their mind and body, in both operation and reception, should also be seen as live, with the participants “being there.”

Summary: Performance Animation

In summary, animation and puppetry have intricate convolution as shown in the following:

- *Ontological-Phenomenological Similarities*: Stop-motion shadow-puppet animation shares with theatrical shadow-puppet shows the physicality of puppets and projection of images;
- *Operation-Reception Synchronicity*: Real-time performance animation in digital media differs from physical puppet shows in materiality, but both have synchronicity of producer operation and audience reception.
- *Puppeteerly versus Viewerly*: Real-time systems turn viewers into puppeteers, animators into audiences, animation into live performance, which I call “puppeteerly animation” as opposed to the traditional film-based “viewerly animation.”

Puppeteerly animation suggests a new notion of live performance in which spatiotemporal co-presence of performers and audience is not required. This new meaning of “live” also characterizes generative and interactive animation. In the section “The Performer” of the next chapter, I will define liveness of interactive generative animation.

The text of the section “The Image” of this chapter, in part, contains material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2009a, 2009b). The dissertation author was a co-author of this paper.

CHAPTER 4

NEW PRINCIPLES OF INTERACTIVE GENERATIVE ANIMATION

Generative and interactive animation, which I argue requires a new paradigm for analysis and design, gives viewers or users an embodied experience of an expanded illusion of life. Today's multimedia artifacts are reactive, autonomous, evolving, and contingent. The previous chapter delineated the four areas of knowledge to provide theoretical underpinnings. That framework corresponds to the four constitutive parts of the emerging phenomenon, namely the observer, the image, the body, and the performer. An animated artifact must be manifested materially in the form of an image and perceived by an observer. The artifact engages the observer in sensorimotor spatiotemporal experience through her or his body. The observer interacts with the artifact and becomes a performer in a live performance (Figure 4.1).

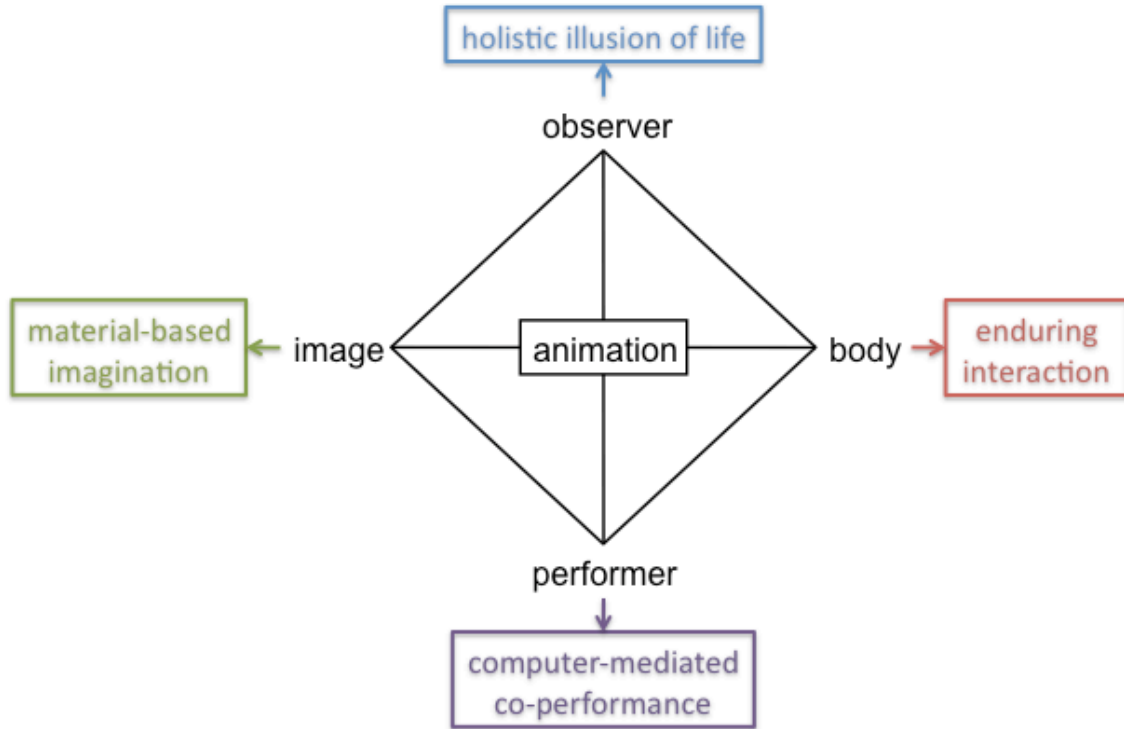


Figure 4.1 The diagrammatic relationship of four constitutive parts of the new animation paradigm

Each of these four integral areas brings about new principles of the new animation paradigm. These principles envelop the characteristics of generative and interactive animation in the digital age:

- The observer perceives a *holistic illusion of life*.
- The image invokes *material-based imagination*.
- The body is engaged in *enduring interaction*.
- The performer, including the observer and the artifact, participates in a *computer-mediated co-performance*.

These principles underscore human perception, cognition, and bodily interaction in connection with animated artifacts. Readers sensitive to animation might notice that the first two principles also apply to traditional filmic forms of animation. They can be regarded as additional tools for rethinking and analyzing animation film. Meanwhile, the latter two principles focus on viewers' bodily engagement and participation, which are

more specific to digital media artifacts. In other words, the sum of these principles characterizes generative and interactive animation, however they are also able to shed new light on traditional principles of animation.

This chapter describes each principle accordingly, with emphasis on how generative and interactive animation exhibits all these exceptional qualities. In order to cautiously observe, identify, and compare the qualities of an artifact with one another, I devise qualitative variables to articulate each of these attributes. At the end of each section, brief sample analyses are included to illustrate the meaning of the corresponding variable, together with its possible values. (Detailed analyses will be covered in Chapter 6). These variables jointly describe a new taxonomy of generative and interactive animation, as well as animated and multimedia artifacts at large. This taxonomy, instead of being a formal classification, suggests an analysis and design paradigm and perspective for the creator who is to design, contrive, and develop multimedia artifacts of this emergent kind. (Chapter 5 will talk more about this taxonomy and Chapter 7 will describe the new design perspective).

The Observer: Holistic Illusion of Life

In the section “The Observer” of the preceding chapter about perception of animacy, I have argued that rather than focusing only on degrees of liveliness, we should pay more attention to the nuances in different kinds of liveliness, like those between prominent behaviors and emergent patterns. This section further shows that how different types of liveliness are simultaneously exhibited in generative and interactive animation, which thus manifests a holistic sense of liveliness. Meanwhile, this holistic illusion of life also cast doubt on the long-standing notion of organism versus machine, and then the conventional animator-centric and program-driven paradigms of animation production. This provocative view characterizes an indispensable part of the new animation paradigm.

Organism versus Machine

In the section about perception of animacy, I briefly described the subtle connection between the animistic and mechanist views of life. This intriguing relationship lies in the essence of the notion of animation. As Cholodenko puts it, the double meanings of animation, endowing with life and endowing with movement, destabilize the opposition of animism and mechanism. He points out that the animatic apparatus animistically gives life to inanimate images by means of mechanical movement driven by viewers’ hands (Cholodenko, 2007, pp. 495-497). The coupling between the apparatus and mechanical movement parallels that between the computer and procedures. However, Cholodenko only touches upon the impact of computing technology on animated film images in terms of visual effects. I would add that generative and interactive animation casts even more disturbances on the dichotomy of organism and machine.

As I have mentioned earlier in the section about perception of animacy, the animistic view of life hinges on certain spiritual and preconscious connection between life and dynamic phenomena. Anything that shows dynamics is seen as having life and soul. In contrast, the mechanistic view, arguably emerging since the Enlightenment in Euro-American thought traditions, aims to dismiss the spiritual connection. People supporting this view are interested in implementation of life-like phenomena by mechanical means so as to demystify life and soul. Yet, at the outset they were only able to implement simple rhythmic and repeatable phenomena, like orbiting, up and down, swinging, waving, and so on, in artifacts like mechanical clocks and automata based on geometry of gears and bars. This “movement” has brought about a contending relationship between organic phenomena and mechanical motion, with the presumption that the latter seems unnatural and uncanny relative to the former. In other words, the inherent type of motion generated by machines is regular and boring.

Natural and Organic Phenomena

With the advent of the computer, more complex dynamic motion, such as an automobile drifting or projectiles being thrown, can be computed and rendered in programs according to physical models. The use of machines to implement motion has been extended to cover physical phenomena. However, it was not until the application of pseudorandomness, such as the Perlin noise function, in computational media that machines could really step out of regularity. Natural phenomena, such as raining, snowing, lightning, leaves falling, splashing water, and the like, were considered so stochastic that could not be easily approximated yesterday, but now they can be realized by various heuristic algorithms incorporating an appropriate component of noise, as seen in many multimedia websites or CGI movies. Organic motion that is pervasive in natural environments also has replicates in the digital domain. For example, Craig Reynold’s computational model *Boids* (1986) simulates the complex and emergent behaviors like

flocking of birds or schooling of fish. The model takes into account three different steering mechanisms, namely to avoid crowding local flock-mates (separation), to align with the average heading of local flock-mates (alignment), and to keep close to the center of local flock-mates (cohesion) (Figure 4.2). It suggests the framework of computer-generated crowd and flock animation, not just constituting major visual effects in many movies but also providing life-like ambience in digital environments, such as the Apple iPhone entertainment application *Koi Pond* (Figure 4.3). All these dynamic phenomena in digital environments are part of the paradigm of generative and interactive animation, which demonstrates its significance in presenting both interactional and emergent liveliness.

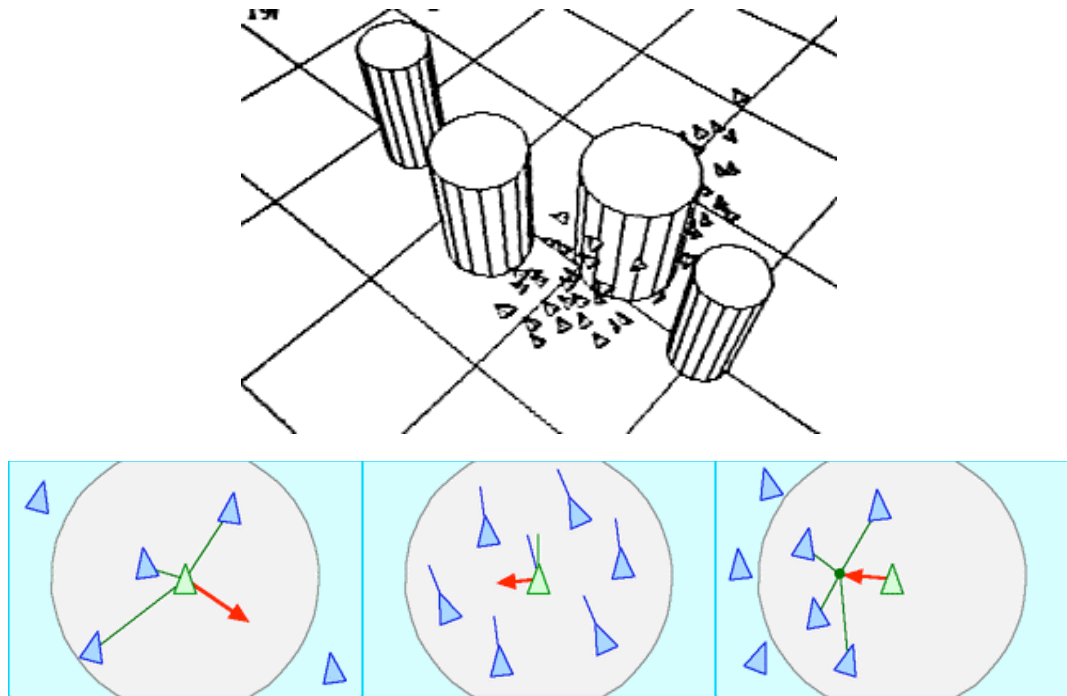


Figure 4.2 Bird flocking simulated by Craig Reynold's *Boids*, modeling different steering behaviors: separation, alignment, and cohesion

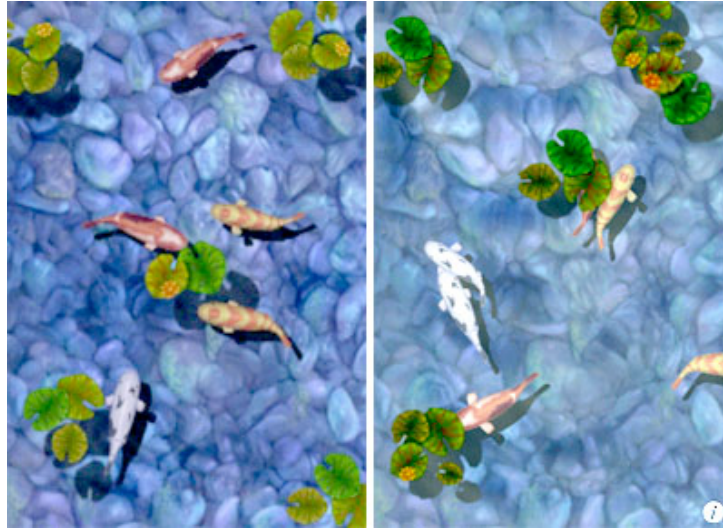


Figure 4.3 Snapshots of the iPhone application *Koi Pond*

Autonomous Behaviors

The distinction of organism from machine is further challenged when some more sophisticated organic behaviors, like autonomous action and reaction, can be simulated in digital environments based on techniques like artificial life (AL) systems, constraint-based programming, or other artificial intelligence (AI)-based approaches. For instance, the former Oz project led by Joseph Bates in Carnegie Mellon University had one of its objectives to be building autonomous and interactive characters that make audiences believe they are alive. The team took reference of perspectives from animation artists, mainly Disney and Warner Brothers' giants, to centralize emotion in creating their autonomous characters. Each character had its emotional state that combined with its personality, influenced how it acts and behaves (Bates, 1994). To me, the project interrogated not only what I meant by liveliness but also aliveness. As Bates put it, they were concerned about "degrees of believability," which depended on not just reactivity but also appearance of goals, emotions, social interactions, and others. Yet, an incident mentioned by Bates resonates my notion of intentional liveliness. A particular character by mistake kept repeating an apparently foolish action, which to their surprise projects observers a personality and so an illusion of life (Bates, 1994). This kind of irregular

occurrence of programmed behaviors caught viewers' attentions and provoked the intentional stance, giving rise to intentional liveliness.

Growth and Metamorphosis

Other essential phenomena of life also include growth and metamorphosis, which can also be modeled computationally. For instance, the growth processes of many species of plants can be reproduced by means of many variants of the L-system based on formal grammar, of which the recursive framework lends itself to various self-similarity forms in the natural organic environment (Figure 4.4). A user would easily control the growth pattern of a tree model in virtual space by adjusting simple parameters. This approach gives rise to myriad fractal-like forms of life in computer graphic content. Besides, a multi-user platform allowing users to submit contents can also project an illusion of self-evolving phenomena. The multimedia website *Ecotonoha* (www.ecotonoha.com), supported by NEC, is an exemplar of this kind (Figure 4.5). The website hosts a virtual tree in which every leaf is a message added by a Web user. The more users leaving notes, the more flourishing the virtual tree is. To a visitor, the virtual tree grows even more vigorously than actual trees.

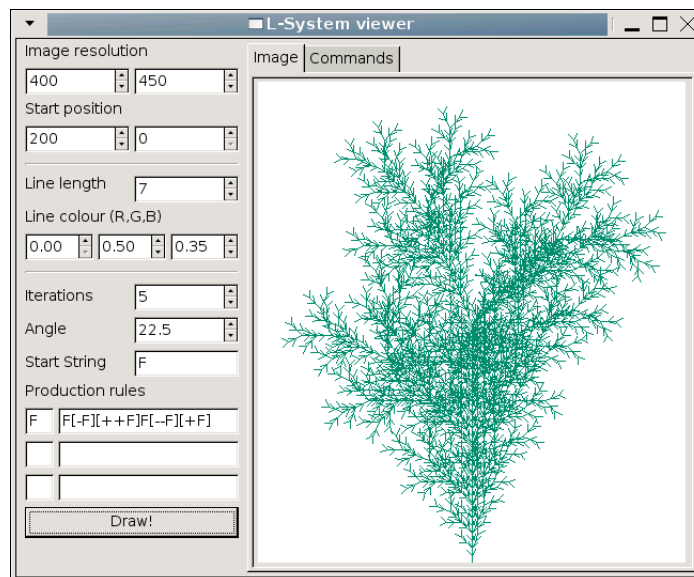


Figure 4.4 A screenshot of an application visualizing the L-system in a tree-like form



Figure 4.5 A snapshot of the virtual tree in the website *Ecotonoha*

The amplification or acceleration of growth and metamorphosis characterizes a pervasive quality of today's generative and interactive animation. As the American cinema theorist Vivian Sobchack points out in her edited volume about the phenomenon of digital morphing, morphing is uncanny, and “strangely familiar,” because of its “‘supernatural’ acceleration of the process of change and transformation” and “effortless accomplishment” in digital culture and economy (Sobchack, 2000, p. xii). Later in her own essay, Sobchack further remarks that morphing, compared with Soviet montage and *mise-en-scène*, presents in cinema a new “reversible temporality” (Sobchack, 2000, p. 137). She raises the “liquid metal” robot in *Terminator 2* (1991) and the malleable bounded city in *Dark City* (1998) to illustrate how reversible and effortless transformation engages audiences like the “playful cartoon physics of animated films.” This notion of transformation echoes what I have been arguing the emergent liveliness, which arises from rhythmic or stochastic patterns that are not obviously progressive. Digital morphing, together with other non-progressive phenomena like snowing and flocking, constitute an array of emergent liveliness manifested by generative and interactive animation.

Animator versus Program

The aforementioned various animation phenomena show that today's digital media are imbued with illusions of life, in which both emergent patterns like swarms or metamorphosis, and prominent behaviors like chasing have been manifested in terms of generative and interactive animation. I call this property a holistic illusion of life, for it encompasses both sides of liveliness.

However, these animation phenomena supported by computation still seem too mechanical and programmed to many animators. Instead, they believe that the act of animation is sensual. They insist that animation must be done by hand, whether it is by drawing in cel animation, by posing a puppet in stop-motion animation, or by keying major frames in computer-aided animation. As Paul Wells and others put it, the first acts of animistic animation come from the automatic and preconscious coordination of the eyes and the hand of the animator (Wells, 1998, p. 32). In the case of traditional hand-drawn animation, the animator draws by one hand and flips a few pages by the other, observes the animated movement, makes necessary changes and flips again. The final animated movement is an outcome of the animator's sensual eye-hand coordination. Therefore, many animators are skeptical about the sense of liveliness implemented by programs. It is the reason why character animation, even for CGI characters such as Buzz Lightyear in *Toy Story* (1995), is also keyed by human animators in authoring software.

This animator-centric view has also been challenged by today's interactive generative animation systems. First, as I have pointed out, the rhythmic and chaotic patterns of liveliness can be generated by pseudorandom functions, such as the Perlin noise function. The inclusion of noise makes the autonomous output of a generative system look less regular, less predictable, and so more life-like. Second, some Disney's animation principles, like follow-through secondary motion, or even walk cycles, can be automated in software of today. The effect of follow-through motion is not simulated by physical models, but rather "mimicked" by mathematical sequences like frame offsetting,

which can generate a rhythmic motion of a ponytail indiscernible from those drawn by hand. Yet, these rhythmic and chaotic patterns are still deterministic to designers, because they can always be reproduced. In order to make the system even livelier, even unanticipated by the designers, the system has to be interactive, which means direct motor control with immediate sensory feedback. To an author using a software application, as well as a participating observer tinkering with an interactive program, the system mobilizes an immediate and automatic motor-sensory coordination. The author or observer is able to tune the motor input in real time in response to the animated feedback and takes the result. The outcome is also based on one's eye-hand coordination. In other words, the preconscious and sensual condition insisted by the animator-centric view is also facilitated in interactive generative animation systems.

An exemplar of this kind is Iwai Toshio's sandbox video game *Electroplankton* for Nintendo DS gaming device. The program generates plankton moving autonomously. A player can use a stylus to interact with the animated plankton to generate spontaneous music, which in turn feeds back to the system for generating succeeding motion (Figure 4.6). This creation process is not unlike abstract surrealists' automatic drawing. It is also like Len Lye and other experimental animators' "direct film" technique (animation directly drawn on film celluloid), but only that the motor-sensory loop in the new system is much more transient and instantaneous. The immediacy and transience of interactive generative animation foreground the polymorphic (in the biological sense, different forms in the life cycle) feature of computer systems as perceptual metamorphic movements, and behaviors, which constitute an expanded illusion of life, including the progressive, prominent "foreground" and the reversible, emergent "ambience." It not only animates the image; it "animates" the computer and destabilizes the dichotomy of the animistic and the mechanistic views, further questioning the separation of organism and machine.



Figure 4.6 Interacting with *Electroplankton* using a stylus

Summary: Toward a Holistic Illusion of Life

In summary, generative and interactive animation exhibits all-round liveliness in terms of the following different forms of computer-generated animation:

- Interactional liveliness can be found in physical effect animation and constraint-based animation.
- Intentional liveliness is shown in AI-based animation.
- Emergent liveliness is demonstrated in digital morphing, crowd animation, and natural effect animation.

I believe that the conventional folk classification of animation phenomena into animistic/mechanistic, organic/mechanical, animatic-centric/program-driven has become ambiguous, because these simple distinctions, as I have argued, are reductive when animation is expanded by generative and interactive systems. My perspective on animation phenomena as prominent behaviors and emergent patterns rooted in perception of animacy does not presume an intrinsic distinction between organism and machine, but rather suggests a typology of liveliness that spans a holistic illusion of life, from the high-pass, progressive, and prominent “foreground” that draws observers’ specific attentions, to the low-pass, reversible, and emergent “ambience” that immerses observers in a life-

like atmosphere. Both sides of liveliness have been actively “enlivened” by the generative and interactive form of animation.

Illustrative Analyses: Types of Liveliness

Analyzing the liveliness of a digital media artifact needs consideration of both sides: (1) any prominent motion showing dynamic tension or invoking the intentional stance; and (2) any emergent pattern that keeps changing but makes no irreversible consequence to the whole environment.

Prominent Behaviors: Intentional Liveliness

In today’s prevalent GUIs, many icons are animated to show prominent motion. For instance, in the Macintosh OS X environment, an icon whose application is not in focus might suddenly pop up and down in the dock in case that it has a message or request for the user (Figure 4.7). This unexpected, but eye-catching, action looks like someone jumping vigorously so as to draw attention, inciting the user to take the intentional stance, thinking about the reason behind this act. The GUI shows intentional liveliness here.



Figure 4.7 An icon popping up in the dock of Mac OS X

Moreover, if the user clicks on an icon of a minimized window, the window stretches and twists into place like a playful genie coming up surprisingly from its lamp. Recall the first time someone saw this effect in the interface. One cannot help wondering why the window has to enact this “bodily performance.” It seems that it “wants” to

differentiate its capabilities from previous conventional windows. Again, this effect presents an example of intentional liveliness.

Animation showing intentional liveliness can also be seen in physical communal environments. For example, the traffic lights mentioned earlier in Chapter 2 display LED-lit animated “green men” to instruct pedestrians when and how to cross the street. During each green signal, the little green man walks slowly, and then starts running faster and faster. The change in speed not only draws pedestrians’ attentions, but also invokes the intentional stance. The liveliness here implies that dangers may be approaching.

Prominent Behaviors: Interactional Liveliness

Another type of prominent behavior is interactional liveliness, which is pervasive in video games. One of the earliest mainstream examples, as I have mentioned, is *Pac-Man*. Apart from a few other programmed behaviors, ghosts in the game are able to switch between approaching and escaping in response to Pac-Man’s power status. The interchangeable predator-prey relation shows dynamic and complex interaction. After that, innumerable combat games (e.g., the *Street Fighter* series) have followed this principle of conditional behaviors with increasing complexity in rules and constraints, immersing us in a substantive visual experience of this kind of liveliness.

Emergent Liveliness

On the other hand, many computer systems exhibit another side of liveliness: emergent pattern. Following the tradition of non-narrative films like Oskar Fischinger's *Composition in Blue*, John Whitney's *Arabesque*, Erica Russell's *Feet of Song*, and others, generative and interactive animation presents the liveliness of rhythm and dance. One canonical example is John Conway’s *Game of Life* (1970). The computer program incorporates a set of transformation rules for 2-dimensional patterns and applies to a grid of pixels. Running the program results in a visual display of a continuous and persistent

change in pixel patterns like choreography of pixels (Figure 4.8). This kind of transformation, in Sobchack's words, is phenomenologically and temporally reversible.

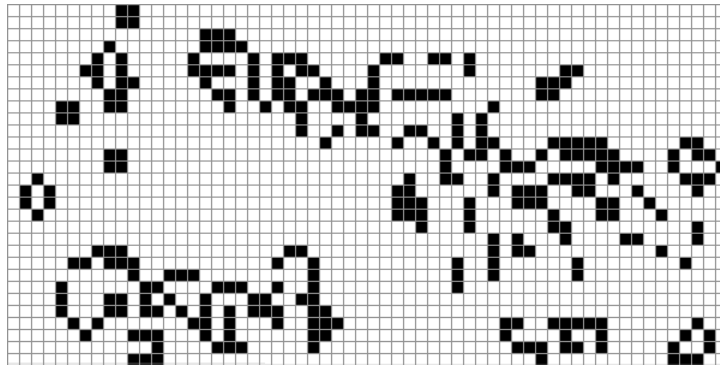


Figure 4.8 A snapshot of an execution of *Game of Life*

The computer capability of creating emergent patterns engenders countless so-called “ambient animations” for screensaver programs, video backdrops, motion graphics sequence, or even movie titles. A latest exemplar is the aforementioned Nintendo DS sandbox video game *Electroplankton*. The game features a milieu with various types of plankton composing several mini-games. Each mini-game allows players to interact with the plankton, which keep making sound and repeating certain action accordingly. The outcomes are like musical and ballet performances of plankton orchestrated by the players. The sound and motion in these mini-games, although each has its own mechanics, share one fundamental principle: looping. When the loop continues, a player is able to tinker with the plankton in order to vary the output. The resulting animation shows no pre-defined beginning and ending, but just lively emergent patterns (Figure 4.9).

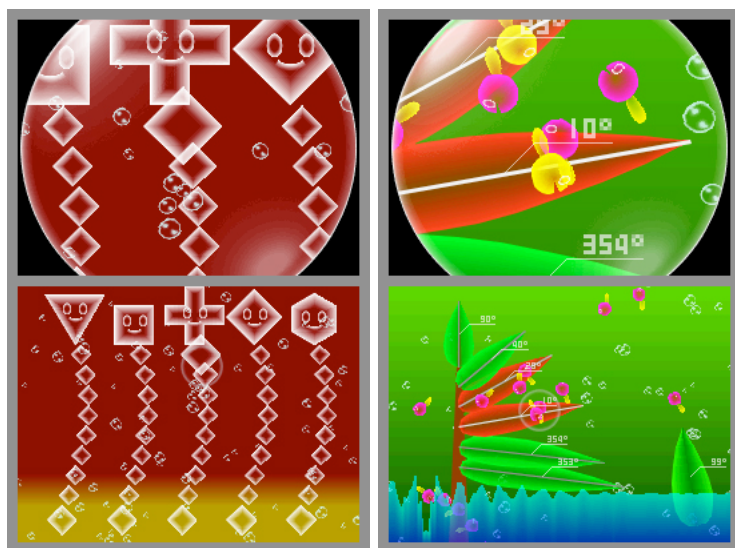


Figure 4.9 Screenshots in *Electrop plankton*

The Image: Material-Based Imagination

In the section “The Image” in the preceding chapter, I argued that animated visual images result in an embodied kind of cognitive processes. In this section, I continue describing how animated visual images mobilize a reflective cognitive process in which material-based imaginative construction and elaboration can take place through conceptual blends. I call this process as “material-based imagination,” characterizing another quality of generative and interactive animation. This kind of imagination is pervasive in today’s digitally mediated environments.

The section about understanding images reviewed major thoughts related to embodied and distributed views of cognition, and their contributions to the link between human perception, cognition, and bodily interaction. Conceptual blending theory is important for marking the link from image to imagination, because it introduces the emergent structure of integration networks that allows imaginative elaboration. Material anchors for conceptual blends provide a good model for integration of the material and the mental, suggesting that perceptual patterns can constitute part of human cognition. Glenberg’s notion of mesh is useful in explaining the matching of perceptual patterns and bodily experiences, which directly addresses embodied conceptualization. Therefore, material anchors and meshes together inform the cognitive grounding of perception and bodily interaction. However, meshes do not have the generative or elaborative power of blends. For Glenberg, patterns projected from experiences cannot be distorted. They must be “clamped,” because loosing the tie for prediction can dangerously result in hallucination (Glenberg, 1997). Moreover, as I pointed out in Chapter 3, Hutchins’s material anchors largely address stable images. In other words, both material anchors and meshes emphasize stability and resist imagination. Hiraga’s iconicity also focuses on the relationship between percepts and concepts and furthermore provides the flexibility of physical-mental mapping that Hutchins’s material anchors and Glenberg’s meshes lack,

yet her work only concentrates on static components like shape and spatial composition but ignores movement. All in all, none of these models are perfect for describing the embodied understanding of animation.

In fact, the new animation paradigm I have been arguing for marks the kind of material artifacts that have more flexible and adaptable qualities for imaginative mental constructions. These artifacts have distilled visual forms and compositions, like Chinese ideographs, architectural sketches, or film storyboards. They have actions and movements carrying visceral sensations, like shadow plays or animated image sequences. All these images not only invite people to perceive but also trigger producers to take further action. This kind of images always constitutes a successive “bouncing” process, in terms of blends, between material artifacts and mental images in the producer’s mind. The bouncing “speed” may vary, from the delayed – in the case of creating storyboards, rough cuts, or flipbook animation, to the instantaneous – in the case of shadow plays or real-time interactive animation. In all cases, blends of material and mental images allow the bouncing process to elaborate successive imaginative constructions. I call this imagination “material-based” based on how one perceives and acts upon material images, as opposed to the mentalist perspective of general thought of imagination as a purely mental activity.

Elastic Anchors

As mentioned earlier, Hutchins coined the term “material anchor” to mean those material objects or images with stable patterns and structures “locking down” specific information or constraints for mental operations, whereas I believe that animated visual images “hold” not only information, but also sensation or meaning, which trigger imaginative elaboration in conceptual blending. These animated images are “elastic” material anchors for blends, which hold sensation in place, but not in shape. I use the term “elastic anchor” to describe these imagination-provoking artifacts. In the tradition of

Hutchins's instrumentally oriented examples, material images and mental images are largely equivalent. The outputs of these blends, like the timepiece example, are fairly fixed as an entrenched cultural model (Fauconnier & Turner, 2002; Hutchins, 2005). For elastic anchors, the subtle difference between the visceral sensation represented by the animated image and the mental image engenders nuanced imagination. A product designer needs to sketch out different views of his or her design (a mental image) on paper in order to further develop the idea after seeing the sketches (a material image). A calligrapher practicing Chinese calligraphy has to repeat writing and looking at the words, making continuous assessments and adjustments. Practicing animators are also well aware of this kind of iterative processes. For example, an animatic sequence might suggest a visceral sensation, with which the animator can compare an intended mental image, combining partial structures from the animated image and the mental image respectively to form a new imaginative image, and then triggering adjustment or modification to the material image. This reflective process iterates and ultimately approaches the imaginative interplay of the material image and the mental image.

Such nuanced interplay would go even more unnoticed when the reflective process becomes instantaneous. It can be illustrated by the case of shadow play. Consider the difference of natural moving images (e.g., incidental shadows) and author-intended animated images (e.g., shadow puppetry) (Figure 4.10). The two images can be materially the same. To Hutchins, neither representation may be valuable as a material anchor because they are not faithful representations of an object, they are just silhouettes. However, the latter can be an elastic anchor for conceptual blends in which the silhouette in action embodying a visceral sensation is blended with the viewer's mental image of an entity (whether human, animal, or even an anthropomorphic object) moving in a similar fashion, thus forming an imaginative understanding that the shadow is cast by an actual character in that mood. On the audience side, this blended image is the meaning of the puppet show. On the puppeteer side, this might be an interim image with which the

puppeteer would fine-tune for another animated shadow. In both cases, the material image blends with the mental image to give rise to next imaginative image.



Figure 4.10 Incidental shadow vs. intended shadow

It follows that material images and mental images have a very intricate relationship. In goal-specific computational operations, they can be regarded as largely equivalent. In creative mental operations, to give an illustrative analogy, they may seem like dancing or boxing partners irregularly approaching each other, whether collaboratively or oppositionally. Moving images can be a vehicle for reconciling both understandings of this intricate relationship, because they constitute a specific type of embodied cognition process. Animated images, with their distilled visual forms, evocative movements, and the material-based reflective process, serve as an excellent elastic input for conceptual blends, because the flexibility and compatibility facilitate the partial structural projection between two images that gives rise to new blends and imaginative images.

Elastic Coupling of Animated Images and Motor Action

If the animated images are instantaneously reactive, as in shadow play, the elastic anchors include not only sensory perception, but also motor action. In digital interactive media, elastic anchors are even more adaptable because a user might interpret his or her

motor input quite variably according to the perceived generative feedback. For an interactive system, generative animated feedback often defines, or re-defines, the meaning of motor action. For example, when a user scrolls the view of a window using two fingers on the touchpad of an Apple MacBook laptop, the scrolling effect defines the action as moving the viewpoint because the scrolling direction is the same as the finger motion. When a user touches and moves a finger on an Apple iPhone's screen, the contrasting scrolling effect (just in the opposite direction) re-defines the action as moving the panel instead. A few other examples include leaving marks on a touch screen defines the touch-and-move action as drawing, a swarm of particles following a pointer defines the action of moving the pointing device as choreographing, turning a new page on the iPhone's screen defines shaking the cell phone as flipping, the magnification effect of the Dock in Macintosh OS X defines the mouse-over action as considering, and giggling human figures in a viral interactive advertisement defines mouse-over as tickling. As mentioned in the section "The Body" of the preceding chapter, what Merleau-Ponty calls "the intentional arc" works beneath perception and absorbs motor action as bodily knowledge. Motion-based motor action embodies our consciousness toward something. Likewise, in digital interactive media, motion-based motor input embodies users' consciousness "moving" toward animated feedback. Since the animated feedback is generative, programmable, and variable, the embodied meaning of motor action becomes adaptable. It follows that a coupling of animated visual images and motor input may yield adaptive and evocative meaning through imaginative blends. Consider the magnification effect of the Macintosh OS X's Dock, both animation (the magnification) and action (mouse-over) may seem non-representational, but the coupling is meaningful when it blends with some everyday experiences (e.g., an individual asserting "pick me!" from a line of candidates).

This coupling idea echoes what I mentioned the motor-sensory feedback loop that characterizes generative and interactive animation. Motor action triggers animation that

in turn incites further bodily engagement. This loop sometimes is delayed, like in hand-drawn animation pencil test, or computer animation preview that needs rendering. Occasionally, the loop is instantaneous and continuing, like that in shadow play or real-time interactive systems. This loop, which relies on the materiality of images that one perceives and acts upon, is the core of elastic anchors. Elastic anchors are elastic in the sense that the meaning is flexibly dependent on the enduring interaction, which will be discussed in depth in the next section “The Body” of this chapter, between motor input and animated feedback.

Summary: Properties of Elastic Anchors

In summary, elastic anchors are characterized by the following properties:

- *Material-based Imagining*: They are material images, consisting of imagic and/or diagrammatic qualities.
- *Imagination Triggering*: They “hold” information or sensation, which can be reflected back to perceivers, yielding imaginative blended images.
- *Action Inviting*: They “invite” perceivers to take motor-action, like modification or interaction.
- *Motor-sensory Connecting*: They constitute iterative motor-sensory feedback loops, such as those in sketching of architectural design, pencil test of hand-drawn animation, writing exercise of Chinese calligraphy, shadow plays, operation of the zoetrope, real-time animation preview, and so on.
- *Spatiotemporal Patterning*: When the loop runs spontaneously and continuously, as in animated images, they provide not only spatial and structural patterns, but also temporal patterns (More descriptions about spatiotemporal embodiment will follow in the third section “The Body” of this chapter).

The loop links sensory perception and motor action on material images, forming a generic mental space for an imaginative integration network, which can be called

material-based imagination. This kind of imagination is pervasive in today's digitally mediated environments.

Illustrative Analyses: Levels of Understanding

The analysis of animation-driven imagination in digital media in terms of elastic anchors is primarily based on conceptual blending theory. Here, I would like to give a brief summary of major terminology to be used in following analyses. These terms are just part of specific vocabulary used by Fauconnier and Turner in their book *The Way We Think*. I highlight them because of their high relevance in my analyses of animated images. More rigorous definitions of a full glossary can always be found in many other volumes of relevant literature in the field.

- *Mental spaces* are small conceptual packets constructed for local understanding and action (Fauconnier & Turner, 2002, p. 40), such as a particular scenario of transaction. They contain elements, like who, when, where, and what, and are typically structured by frames.
- *Input spaces* are input mental spaces for conceptual blends. Early examples in blending theory usually involved two input spaces.
- *Cross-space mapping* are partial connections between counterparts in the two input spaces in a blend (Fauconnier & Turner, 2002, p. 41).
- *Selective projection*: Elements and relations from the inputs are only selectively projected to the blend.
- *Compression*: Cross-space mapping, or called "outer-space" links between the inputs to the blend, will be compressed into relations inside the blend, or called "inner-space" relations (Fauconnier & Turner, 2002, p. 92).
- *Vital relations* are important recurring relations in compression (Fauconnier & Turner, 2002, p. 92).

- *Frames* are long-term schematic knowledge (Fauconnier & Turner, 2002, p. 40), such as the sell-buy frame in the modern society, which includes roles for buyer and seller, merchandise, price, transaction period, terms of contracts, and so on.

The following is an example mentioned by the dissertation author and D. Fox Harrell elsewhere (Chow & Harrell, 2009b). Consider the “water-level” battery meter interface of the Japanese mobile phone N702iS (Figure. 4.11). The phone was manufactured by NEC Corporation in 2006. It comes with an intriguing gestural interface designed by Oki Sato and Takaya Fukumoto. The interface displays images visually resembling liquid water. The subtle movement of water in the interface invites the user to tilt the cell phone a bit, resulting in further reactive movement of the computer-generated water images. This motor-sensory loop runs spontaneously and continuously, building a material-based mental space and holding an embodied meaning (the action and reaction afforded by this artifact). The reactive image anchors this meaning as a direct input to the blend with the user’s sensorimotor experience of tilting a bottle of water (similar to meshes of percepts with bodily experiences), yielding an imaginative conceptualization of a cell phone containing water inside, in which elements including animation (flowing), action (tilting), and artifacts (the phone and the bottle) are selectively projected and compressed into the blend. These compressions rely on a match of motor-sensory patterns in the immediate situation and from memory, but the patterns encode not only spatial information (where to grip, which direction to tilt, etc.) but also temporal data (how fast to move, when to pause, etc.). These are motion-based input⁶. Imagine how, if the water image did not flow to the tilted side, or it flowed with a delay of several seconds, the match and the blend would be rendered void. Hence, in the blends involving elastic anchors, the spontaneous and continuous loop is the core.

⁶ The section “The Body” of the preceding chapter has described this kind of input, whose motion qualities are significant to output.



Figure 4.11 The interface of the mobile phone N702iS showing computer-generated imagery of water reactive to user action

The following diagrams illustrate two levels of blends emerging in the use of this artifact. Based on Fauconnier and Turner's traditions (Fauconnier & Turner, 2002), an integration diagram depicts the input spaces horizontally, with the output of the blend in the bottom. The lines between these spaces are mappings, selective projections, or compressions of vital relations. The top structure is the generic space, which encapsulates elements shared by all input spaces. In this structure, I add a circular loop to represent the aforementioned motor-sensory loop. Moreover, motor actions and perceptions are represented in terms of mathematical relations. For example, flow-inside (water, bottle) means that liquid water flows inside the bottle, while tilt (phone) denotes the act of tilting the cell phone.

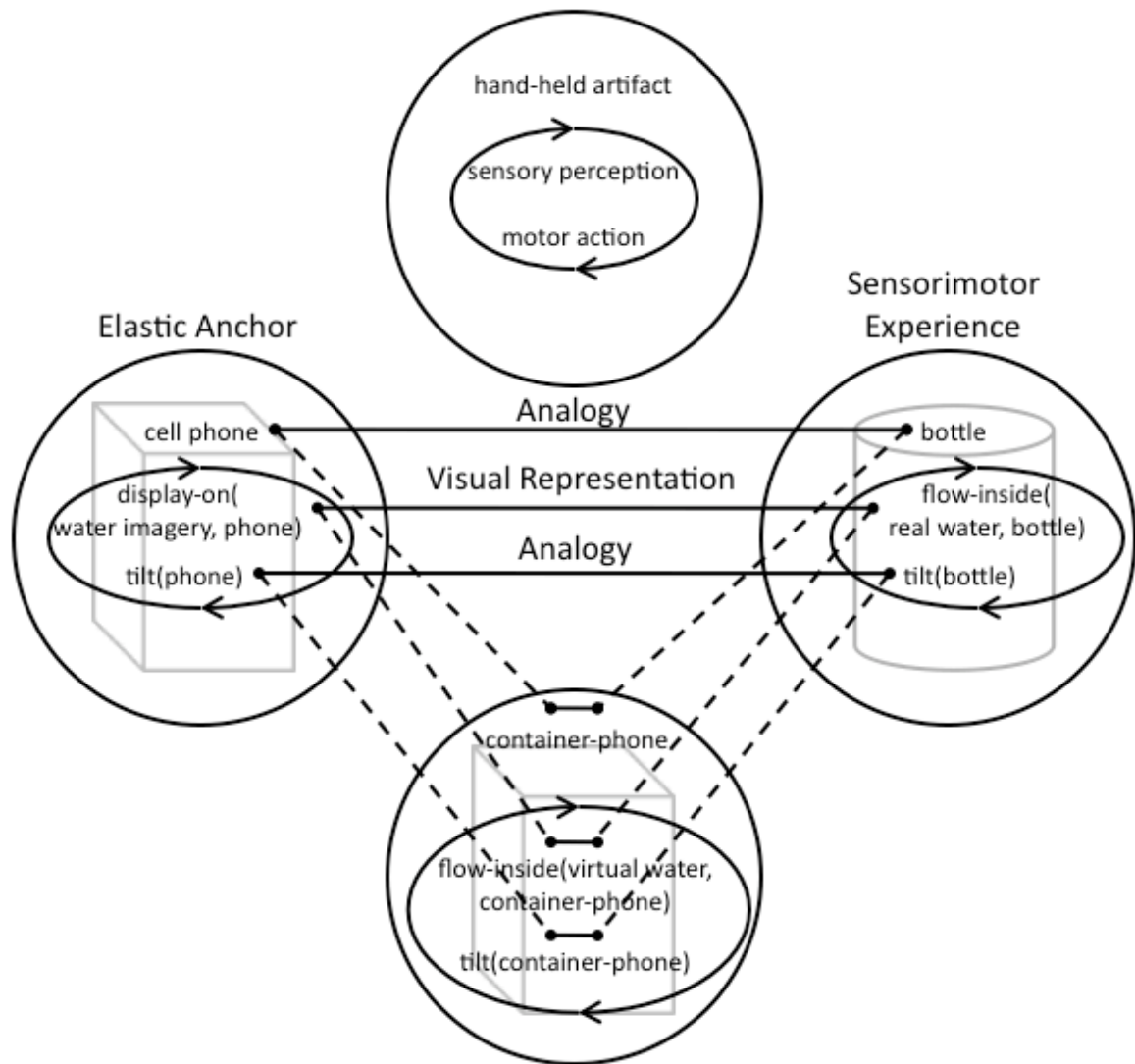


Figure 4.12 The immediate blend of the water-level interface

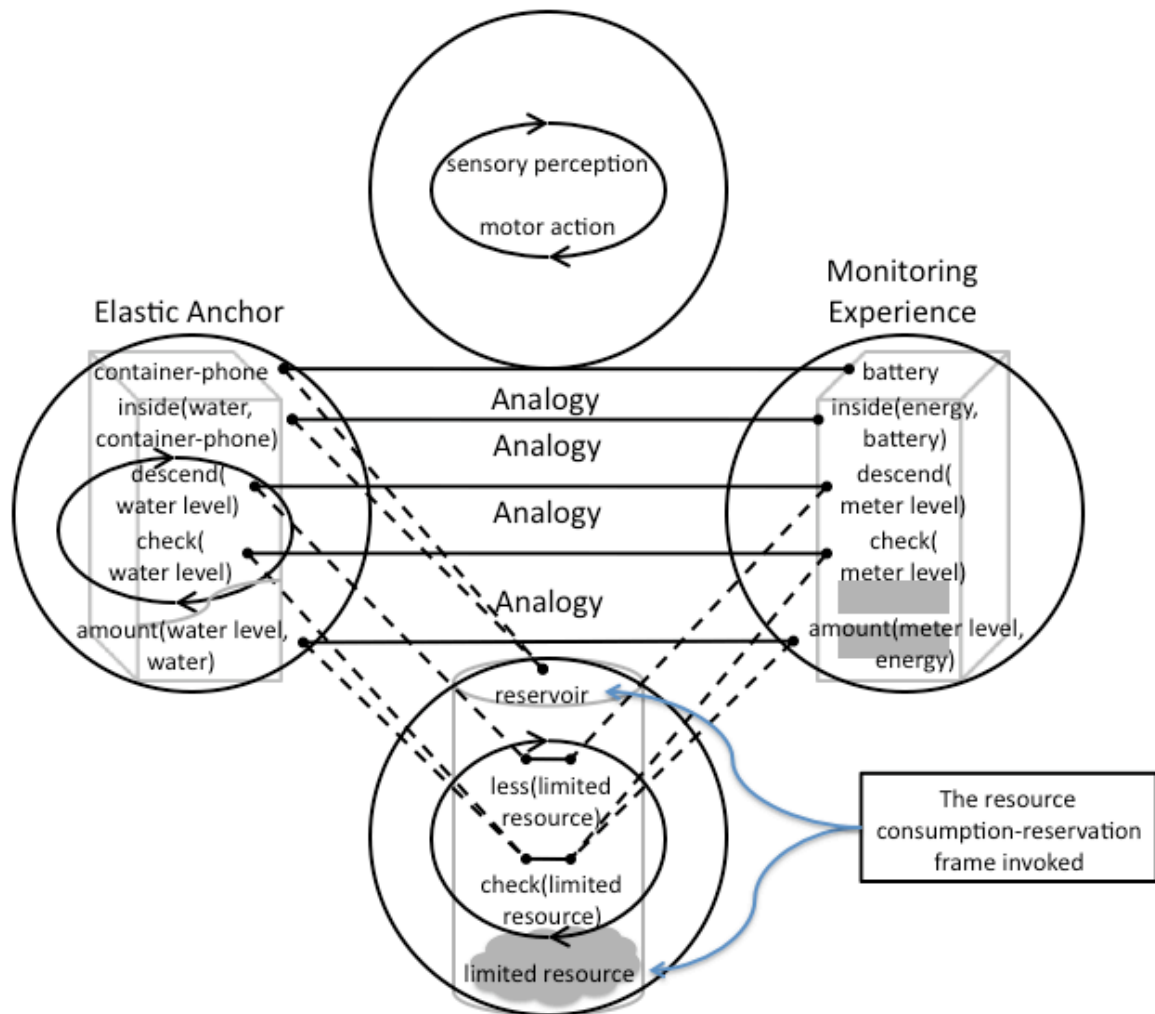


Figure 4.13 The metaphorical blend of the water-level interface

In the first level (Figure. 4.12), the generic space consists of a concept of the motor-sensory loop, together with a hand-held artifact. For the blend, there are compressions of outer-space representation links (a type of vital relations, representation of virtual water imagery displayed on the phone for actual liquid water flowing inside the bottle), and compressions of outer-space analogies (e.g., tilting the cell phone cf. tilting the bottle), forming a new concept (the cell phone contains virtual water). The output is an imaginative blend of a water-filled container-phone.

What makes the artifact even more intriguing is that the computer-generated water level is in synchronization with the battery level of the phone. When checking the water level (just like an ordinary user checking the battery level of an ordinary cell

phone) intermittently, one notices that the water level descends gradually and reflects the ongoing energy consumption in the container-phone. This articulation further incites the secondary blend of water with energy in the battery (Figure. 4.13). The generic space now includes a concept of a level-type indicator, which reflects the amount of a particular resource left. In the blend, the outer-space analogy link between the water level and the battery meter level is compressed tightly into an inner-space category relation: the water level is a battery meter. The blend might then be elaborated into an imaginative narrative about someone in some harsh environment checking how much water is left or similar tales of water-conservation. The monitoring of water consumption invokes a common frame nowadays: conservation of energy or other scarce natural resources. The elaboration of such blend leaves the user a strong and provocative message: “save the juice!”

It should be noted that if the water level does not continually react to user action, or if it is not autonomously descending according to power consumption, the message of the interface would be totally different. The first condition of having an immediate experience of continuing reactivity is necessary for the first level blend. I call this immediate level of understanding. The second condition of coupling the water level and energy consumption supports the imagination of a metaphorical narrative in the secondary blend (In Turner’s sense, the narrative is a small spatial story, see (Turner, 1996)). I call this metaphorical level of understanding.

The Body: Enduring Interaction

In the section “The Body” of the preceding chapter, I call for considering motion-based motor input in designing interactive systems, because motion reveals and embodies intention in both space and time. This section expands this idea of spatiotemporal embodiment to propose a continuously changing and varying digital environment supported by generative and interactive animation.

Interaction between humans and computers is often figuratively interpreted as conversation or dialogue between two active participants (Crawford, 2005, p. 29). In this section, I argue against this conventional view because it overlooks an integral part of our everyday lives: spatiotemporal embodiment of meaning. As Merleau-Ponty puts it, “our body inhabits space and time.” The embodiment of meaning depends on both space and time. In digital interactive media, this goal can be achieved by: (1) as I have argued, motion-based motor input; and (2) a continuously changing and diverging environment. I call the paradigm of interaction mechanisms consisting of these two vital components as “enduring interaction,” in contrast to the discrete, conversational types of alternate computer-human interaction, like the classical command-line interface, or the conventional point-and-click widgets found in many GUIs.

Spatial Embodiment: Lakoff & Johnson’s Spatial Metaphors

To Merleau-Ponty, motility reveals our intentions toward something. After Merleau-Ponty, more thinkers have focused on how our bodies interact with worldly situations. Lakoff and Johnson believe that bodily interaction has direct contribution to human cognition. They assert that most of our cognitive models, named image schemas, are comprehended through bodily projection. For example, our bodies are containers taking in food and emitting waste, so we project a container metaphor onto abstract concepts, as when we understand a tribe as living “in” the forest (Lakoff & Johnson,

1999, p. 36). The theory of image schemas is a powerful tool for us to understand how we reason based on spatial logic. However, it also demonstrates humans' over-reliance on spatial relation while downplaying temporal transition. Because time is an abstract concept, we are used to spatializing time metaphorically. When we say the "flow" of time, we may mean seeing time as a moving object (e.g., a river) from a static observer's viewpoint. Meanwhile, we might see future as something fixed "in front of" us. We often think of how many days "to go" before a deadline, just like time is a static point to which we are moving. In short, we conceive of time in a source-path-goal schema in which we situate our body. It seems that our over-reliance upon spatial reasoning is a result of our body's spatial existence in the world. However, our body also inhabits time. As Merleau-Ponty puts it, "I belong to space and time, my body combines with them and includes them" (Merleau-Ponty, 1962, p. 140). The notion of embodiment should include the concept of time.

Temporal Embodiment: Deleuze's Cinematic Philosophy

Gilles Deleuze has attempted to theorize a perspective of time in his two famous volumes about cinema. He points out that our natural perception introduces "halts" or "fixed points" in our everyday lives (Deleuze, 1986, pp. 22-23). We tend to "immobilize" the continuous flow of life, slowing down the intense "flux" of data, in order to think (Colebrook, 2002b, p. 149). Hence, we see time as connection of those fixed viewpoints within some ordered whole. This model is best illustrated by the use of storyboards in filmmaking, and even in interface design. To an interface designer, a storyboard helps her or him understand the connection between input and feedback. Apparently, time is just a sequence of alternating stimuli and responses. In other words, Deleuze agrees that we are used to creating a spatial concept of time through natural perception. However, he also reminds us "cinematographic perception works continuously" (Deleuze, 1986, pp. 22-23). His major theses include the concepts of the "movement image" and the "time

image.” The former refers to the early cinema enabled by the moving camera and Soviet montage, telling us movement is not just translation in space but also related to variation. The time image further presents the continuous variation, a divergent “becoming,” directly by irrational cuts. I believe cinema, with its hundred years of history, has changed our perception of time, and also the concept of time, making it different from the natural “stimulus-response” mechanism (Colebrook, 2002b, p. 149). In digital media, the old perception of time corresponds to the discrete, conversational style of interaction. For instance, the classical command prompt interface introduces “halts,” breaking the engagement into ordered, discrete, and alternate segments of input and feedback. Today’s digital media artifacts should instead present the “becoming” in temporally embodied interaction, engaging users in continuity and variation. The idea of embodiment should be re-considered as bodily engagement in a lively world, which is not only reactive but also autonomous and evolving. In case of an animated artifact, it should be able to project an illusion of a constantly evolving environment meanwhile diverging with respect to a user’s motion-based input.

Enduring Transformation

The essence of temporal embodiment is that of a constantly changing whole. An engaging animated interface should present continuing and differing variation. This enduring transformation can be manifested with generative and interactive animation. With today’s computational technology, generative algorithms are able to render animation transforming indefinitely. In the section about growth and metamorphosis of this chapter, I described transformation with a focus on the temporal and phenomenological reversibility. Here, I elaborate the idea of transformation further to cover the constantly variable and divergent nature. Algorithms can take user input and process it in real time, supporting immediate feedback to the animation. The outcome is a constantly changing image readily transforming in reaction to its viewer. One of the

exemplars is John Conway's *Game of Life*, which has been mentioned in the section "The Observer" of this chapter. The "game" is a two-dimensional cellular automaton in which the cells live, die, and reproduce according to pre-defined rules based on the neighboring cells. The running of the game looks like an animation of vivid patterns morphing constantly. Meanwhile, a viewer or player is always free to interfere the growth of the cells by putting extra cells in the two-dimensional grid. This motor input is motion-sensitive, because to add cells the user just runs the pointer across the grid. At the perceptuomotor level, the piece is a generative and interactive animation. At the metaphorical level, it projects an image of life, which is constantly evolving because of its rules, and is always differing because of its openness to intervention. The interaction between the user's motion-based motor action and the animated images exemplifies a temporal embodiment in a "plane" of "becoming."

Besides, Craig Reynolds' *Boids* is also noteworthy. As I have mentioned in the section "The Observer" of this chapter, it is a computational model simulating emergent motion of animal crowds, such as flocks of birds and schools of fish. The result of the model is usually called crowd or flock animation. The algorithm deals with motion of individual animal in relation to its neighbors. At the level of the individual, it addresses the translation in space. At the level of the crowd, the model presents a collective whole that looks like an amoeba in enduring transformation. It demonstrates what Deleuze means by "movement," which is "not a change in place" but rather "a transformation of the body," "a body being nothing other than its movements" (Colebrook, 2002a, p. 45). If individual animals were not in motion, there is no transforming amoeba. With the development in real-time control technology, this "amoeba" becomes interactive. The iPhone application *Koi Pond* mentioned earlier has implemented the model to render a pond of koi fish each of which is reactive to user taps. In other words, the whole pond is continuously changing on its own and possibly differing with user intervention.

The Enduring Environment

In interface design, one of the primary goals has been to make interfaces more “user-friendly,” or “usable.” Following the views of Ben Shneiderman, Donald Norman, Jakob Nielsen, and others, the terms generally describe a tool that is easy to learn, understand, use, or manipulate. Their ideas have led to the prevalent user-centric approach to interface design. The paradigm emphasizes usability, efficiency, and productivity, rendering computer a synonym of office equipment or business machine. On the other hand, Paul Dourish’s phenomenological approach suggests designing interactive systems as environments engaging users like the everyday world. This world is familiar and approachable to the users. However, grounded in Heidegger’s “Being-in-the-world,” Dourish’s perspective is about a “work-world,” which is mundane and routine. In designing animated systems like multimedia websites, entertainment applications, or digital artworks, I argue for an enduring and lively world, which is more habitual and intimate to humans. This environment engages human users less with objects and space, but more with change and time. According to Deleuze, time is not some ordered sections of stimuli and responses, like conversation, but rather a continuous and divergent “becoming.” Digitally mediated interaction of animated artifacts aiming at more embodied meaning should present this “becoming,” engaging human users in an enduring and lively world, which not only reacts to motion-based input but also presents continuous transformation. This constantly changing environment can be manifested with generative and interactive animation, making computer interaction more embodied and computer interfaces more reminiscent of our everyday experiences. Some good examples include the greeting page *SnowDays* at Popularfront.com (to be discussed in the following section), the water-level interface in mobile phone N702iS (to be discussed in Chapter 6), and lately the mobile application *SunDial*, which displays the continuous movement of the sun to support religious practices of Islamic users (Wyche, Caine, Davison, Arteaga, & Grinter, 2008). They all use animated images that transform

continuously, just like the running of *Game of Life* and *Koi Pond*, projecting users a strong sense of “becoming” and life.



Figure 4.14 A mobile Islamic call to prayer application *Sun Dial*

Summary: Conditions of Enduring Interaction

In summary, the necessary conditions of an enduring interaction mechanism include:

- *Motion-based Motor Input*: The motion qualities of user input must have significant effect on system output;
- *Enduring Environment*: The system still presents continuous transformation even without any user input. On the other hand, it would show differing output in response to varying input.

I believe more computational artifacts with interfaces of this kind will make computers more familiar, intimate, and close to human users.

Illustrative Analyses: Mechanisms of Engagement

Hence, the analysis of bodily engagement with animated artifacts requires a new model derived from spatiotemporal embodiment of meaning, including motion-based motor input and the enduring environment. I introduce a continuum of engaging

mechanisms from the least to the most embodied, accordingly halted engagement, decelerated engagement, and continuing engagement.

Halted Engagement

Halted engagement mainly refers to a kind of conversational interaction with artifacts or machines. The mechanism involves distinct sections of motor input and sensory feedback taking turns alternately. The ordered connection of these sections is the “immobilized” “becoming” of the environment. The use of mechanical typewriters is a typical example. The typewriter responds to each key tap by the typist with a corresponding character strike in sequential order. Any one tap that did not wait for the feedback of the preceding tap would cause a jam. In other words, the input and feedback sections cannot overlap. This mechanical constraint leads to strictly alternate sections, forming the defining characteristic of halted engagement. Moreover, halted engagement involves no motion-based motor input. In typing a text, motion qualities like how hard the finger hit a key or the path reaching a key could have no significant effect on the outlook of the printed text.

The phenomenon of halted engagement seems to be a result of the mechanics of the apparatus, yet halted engagement should not be thought of only limited to mechanical artifacts. In fact, “halted” examples can be found in digital environments as well, including the command-line environment in MS-DOS and the point-and-click mechanism in most graphical user-interfaces. In these cases, a system usually takes a user input and then responds alternately. The system would wait indefinitely for a user input. It shows that halted engagement is a matter of design, not mechanics.

Decelerated Engagement

Decelerated engagement differs from halted in that the artifact considers motion-based motor input. Yet, the interactive engagement would still be slowed down, so

decelerated, in the case of a delayed input. The result is a kind of semi-enduring interaction. Regarding motion-based input, some machine interfaces provide exemplars. As mentioned in Chapter 3, professional VTRs usually come with a jog dial allowing a user to control video playback by rotating the knob. The motion qualities of the dial affect how the medium is presented. A clockwise spin results in fast-forward, whereas a counter-clockwise spin rewinds the tape. The faster it spins, the faster the tape plays. The case is similar to the mechanics of a zoetrope. The direction of rotating the zoetrope determines the direction of movements presented through the slits. These machines entail motion-based input and manifest more embodied interaction.

However, these machine interfaces still have inadequacies, resulting in imperfect enduring interaction. Although motor input in the zoetrope or the jog dial is motion-sensitive, a delay of user input could slow down the whole “becoming” output. In case of the zoetrope, the viewer needs to keep spinning the apparatus in order to see the animated effect. The spinning speed determines what the viewer exactly sees through the slits. If the viewer defers the motor action, the animation would slow down and finally halt. I call this interaction mechanism decelerated engagement. Similar to the halted cases, decelerated interfaces would also wait indefinitely for user input.

In fact, many computer interfaces are moderately “decelerated.” While they are waiting for a user input, they project an illusion of continuous variation through multimedia, like animation or sound, but the whole environment actually makes no advancement. This trick is best demonstrated by some computer games in which even though a player-controlled character stops moving, other animated objects still keep wandering and the music still goes on. It seems that the environment keeps changing; in fact, the animation and music are looped, and the game does not proceed to next levels until the player takes subsequent action.

Continuing Engagement

Lastly, continuing engagement describes those systems showing ceaseless transformation, with or without user input. This kind of engagement is continuing in that the environment would continue changing autonomously even there is no input. Simultaneously, a user input, mainly motion-based, at any time might trigger a particular variation that would carry on. This “becoming” is persistent and divergent. A good analogy is the Japanese tea-serving *karakuri*. On winding up, it paces slowly with a cup of tea approaching the audience. When it bows, the gesture cues the audience to pick up the teacup. If the audience does, it would wait for the return of the cup; otherwise, it would not wait but turn away, and come back after a while. The mechanical doll is following its own rule continuously. Meanwhile, it also reacts to its audience’s timely action and might have a little difference in behavior. It is able to engage its audience in continuing and differing action. In digital media, this kind of continuing engagement is emerging.

A prominent example is the greeting front page *SnowDays* at Popularfront.com. The page displays an outdoor view of snowing somewhere. The falling flakes vary in shape because they are actually other visitors’ submissions. A visitor may create his or her own snowflake by a simple interactive cutting tool and attach a message. Once the visitor submits the flake, it falls in the scene and constitutes part of the “becoming.” The visitor can “catch” any falling flakes and check out the details and messages. Yet one has to take timely action, otherwise the target may fall out of the window frame. Therefore, motion matters in the user input of this system. Furthermore, even though the visitor does nothing but just watching, the background always changes with the time of day, forming another part of the enduring environment.



Figure 4.15 Snapshots of *SnowDays* at Popularfront.com

Other good examples include the water-level interface in the mobile phone N702iS, which has been mentioned in the “The Image” section of this chapter. The interface displays computer-generated images of water that react to user action. When the user tilts the phone, the way and speed of user motion determine how the water flows on the screen, yielding an illusion of a water-filled cell phone. Shaking more vigorously would lead to other effects like turning off the alarm. That means the user input is motion-sensitive. Meanwhile, even though there is no user action, the water level would continue falling very gradually according to the battery consumption. This subtle change reflects an enduring environment. All in all, the reactive and self-evolving water images constitute a “becoming.”

In short, continuing engagement includes both motion-based input and enduring environments. Decelerated engagement involves only the former but lacks the latter. If both vital components are left behind, an interactive system belongs to examples of halted engagement.

The Performer: Computer-Mediated Co-Performance

In the section “The Performer” of the preceding chapter about live performance of animation, I have shown that the emergent animation phenomena in digital media can be seen as a new form of live performance, which I call “puppeteerly animation.” This section further attempts to define liveness of generative and interactive animation. This notion of liveness contributes to an expanded illusion of life in two folds, namely a mutual sense of liveliness between observers and performers, which I have described when talking about performance animation, and more importantly, the contingent nature of life. That means, the convolution of generativity and interactivity results in contingency, a vital part of the new animation paradigm.

Defining Liveness of Animation

Generative and interactive animation is a form of live performance in digital media, because there is a synchronicity of operation and reception enabled by today’s real-time control and rendering technologies. This technological phenomenon problematizes the traditional notion of liveness. It seems that physical presence, proximity, and immediacy are all outmoded requirements. Philip Auslander has made a similar remark on the phenomenal impact of media technologies and mediatization in performance practices. He provocatively claims that there is no ontological difference between mediatized and live performance, so refuting ontology of liveness (Auslander, 1999, p. 50). He asserts that contemporary live performance is not only mass-consumed by a huge audience but also mass-produced and -reproduced from a recurrent text, just like other media content. Good examples include Broadway theaters that are adaptations of movies (e.g., *Billie Elliot*), the pop singer Madonna’s world tour concerts re-manifesting her music videos, and others. In Walter Benjamin’s words, live performance has lost its aura in the age of mediatization. Yet, Steve Dixon reviews Auslander’s claim

with a critical lens, thinking that his view has downplayed the perceptual differences between live and mediatized performance. He argues that audience reception in the theater differs experientially from that in other media. The former is a collective experience while the latter a voyeuristic one (Dixon, 2007, pp. 129-130). However, there is no such disparity between puppetry and animation, as I have described the shadow theater and the shadow film in the section about performance animation. In fact, the reception and perception of animators and audiences in the new animation paradigm have become even more comparable because of real-time technologies. Hence, I am inclined to say that computer-mediated performance animation is live from the perspectives of participants, including both the performer and observer's views.

Meanwhile, in new animation phenomena, generative algorithms support interactive experiences phenomenologically similar to live performance, which can present different instances on every viewing and show responsive variations on audience feedback. These experiences are definitely scarce in the traditional film form of animation. We need to describe this quality in a formal way. Here, Auslander's interpretation of liveness in today's mediatized culture may shed some light on the matter. He asserts that live performance exists only after recording technologies has engendered what we call mediatization. Inspired by Jean Baudrillard's dictum that the "real" is something that can be reproduced, Auslander defines the "live" as "that can be recorded" (Auslander, 1999, p. 51). Although this kind of oppositional definition seems to be paralipsis to some people, it reminds me that real-time technologies differentiate "live" puppeteerly animation from traditional "playback" viewerly animation, just like what recording technologies did to the live and the mediatized. I propose that interactive generative animation should be regarded as "live" in relation to playback of pre-rendered animation. The latter is something already closed, complete, recorded, and materially stored. The former is still open, emergent, subject to change, and to be recorded, such as *Interactive Wall: BIG SHADOW* (an interactive outdoor advertising event featuring

shadow play connecting online web visitors and on-location viewers in real time), Camille Utterback's well-known digital artwork *Text Rain* (more discussions follow in the next section), the virtual Britpop band Gorillaz' live concert, machinima animation, performance in SecondLife, and others. For this kind of "live" puppeteerly animation, I define liveness to be inversely proportional to the amount of stored or pre-rendered components. High-degree liveness corresponds to less pre-rendered visual content for direct playback. Low degree of liveness means a major playback of stored content. In this regard, screening of Oskar Fischinger's abstract films is not live while running of John Conway's *Game of Life* is highly live.



Figure 4.16 BIG SHADOW found in Tokyo



Figure 4.17 Digital artwork *Text Rain*



Figure 4.18 The Gorillaz performance in 2006 Brit Awards

Improvised Co-Performance and Contingency of Life

As I mentioned in the section about performance animation, real-time control and rendering technologies have made possible a mutual sense of liveliness between the performer and the observer, destabilizing their original perspectives. The designer, animator, or viewers of a multimedia artifact enabled by these technologies all participate in a co-performance. They all become “performers” in the “live” show of the artifact augmented with generative and interactive animation. The artifact contributes to different versions of the performance, because generativity supports pseudorandom variation, and interactivity facilitates human intervention. Hence, each presentation is like an improvised co-creation between the participants and the artifact. This kind of co-creation is more improvised than prepared because the designer can never exactly know how or when the participants would take action to interfere the outcome.

The idea of improvisation is often positioned as an alternative to formal, well-prepared, and structured music performance. With an aim to explore possibilities of music improvisation with computers, the trombone player, composer, and scholar George Lewis has constructed the computer-driven interactive music system *Voyager*. Talking about implication of his masterpiece regarding the controversial opposition of “composed” and “improvised” traditions of music making, he emphasizes that to

distinguish from composition, improvisation must entail real-time intervention. An improvising machine should be “open to input, open to contingency” (Lewis, 2000). His assertion also applies to generative and interactive animation, which gives rise to improvised co-performance, bringing openness and contingency to an artifact.

Regarding contingency, Dixon also ties the concept closely to live performance. As he puts it, live performance always carries the possibility of “unexpected” happening (Dixon, 2007, p. 131). In other words, contingency is an integral part of liveness. The degree of liveness in generative and interactive animation marks the contingency of life carried by an animated artifact. The higher the degree of liveness, the less stored content, and the more variation can be generated by algorithms or driven by participants, the more uncertain the outcome will be. In short, liveness brings contingency of life to an animated artifact, providing audiences with a more embodied experience of the expanded illusion of life.

Contingency of New Animation Phenomena

The aforementioned ideas of improvisation and contingency in animated artifacts labeled as liveness also inform a distinctive characteristic of interactive generative animation, and digital media at large. Regarding the essences of digital media, quite a few media theorists have already made well-received conclusions, including Janet Murray’s four essential properties of digital environments (Murray, 1997), Lev Manovich’s principles of new media (Manovich, 2001), Peter Lunenfeld’s aesthetic of “unfinish” in digital media (Lunenfeld, 1999), and others. This section, rather than arguing for another distinguished point of view, attempts to augment the lively discussion by marking the role of contingency in today’s digitally mediated context. As a subgenre of digital media, the new animation paradigm, with its contingent nature destabilizes the respective roles of performers and audiences, animators and viewers, throwing new light on the broader theoretical context of post-structuralism.

Barthes's pronounced dictum "Death of the Author," which is generally thought of as marking the birth of post-structuralism, provokes subversion to the hegemonic authority of the author and turns the reader from a consumer into a producer of meaning (Barthes, 1977a). As he puts it at the end of his essay, "a text is made of multiple writings, drawn from many cultures and entering into mutual relations of dialogue, parody, contestation, but there is one place where this multiplicity is focused and that place is the reader, ... we know that to give writing its future, it is necessary to overthrow the myth: the birth of the reader must be at the cost of death of the Author." In another essay "From Work to Text," he further talks about the consumption of a text, suggesting that we "abolish (or at the very least to diminish) the distance between writing and reading" (Barthes, 1977b). He draws an analogy from the art of contemporary music, in which the performer or interpreter is regarded as the co-author of the score, to advocate reader collaboration and improvisation in writing. The idea is comparable to Umberto Eco's notion of the open work or its subgenre work-in-movement, which refers to the kind of works that are always open for interpretation, intervention, and to be completed (Eco, 1989). He raises a wide array of examples, including James Joyce's *Finnegans Wake*, Alexander Calder's mobile sculptures, abstract expressionist paintings, and even live television broadcasts in which each director is free to choose from a set of cameras in order to complete his or her narrative. However, Eco's view is less radical than Barthes's concept of writerly text, because Eco presumed that the direction of "movement" in a work is still intended, proposed, and so steered by the author (Eco, 1989, p. 19). In fact, Eco's work-in-movement is more similar to Nelson Goodman's notational system of art, like a musical score or a woodblock print, of which the constitutive component is usually prescribed by the author while the contingent component is open to the performer or the printer (Goodman, 1976, pp. 116-118). All in all, these (relatively) contemporary thoughts mark the contingent nature in the world of creativity.

Actually, what Eco and Goodman describe are author-intended rule-based systems with contingent instances reactive to user actions, which also lie at the heart of interactive generative animation. Instead of completely subverting the role of the author as did the writerly text, the new animation paradigm still assumes that the author plans, designs, or programs the expanded illusion of life, while the user and the computational artifact co-enact various versions of the live performance. The degree of liveness is inversely proportional to the amount of pre-rendered or stored material content. Hence, higher degree of liveness indicates less certain output, resulting in more divergent possible outcomes. The notion of liveness I have been arguing for incorporates the idea of improvised co-performance, better describing and capturing the contingent nature of digital media, in which users always experience differently on each system instantiation and outcomes at all times vary with every user intervention.

Summary: Liveness of Animation

In summary, this section defines liveness for analyzing and designing interactive generative animation, followed by its implication on the contingent nature of life manifested in such artifacts:

- *Liveness versus Playback*: The degree of liveness is inversely proportional to the amount of stored components in playback during the improvised presentation of an artifact.
- *Contingency*: Less pre-rendered, pre-captured materials result in higher degree of liveness, in turn invoking an embodied experience of the contingency of life.
- *Post-structuralist Inclination*: Liveness reflects contingency, being a reference indicator of the post-structuralist inclination embedded in the design of an artifact.

Illustrative Analyses: Degrees of Liveness

Liveness of generative and interactive animation brings the contingent nature of life to digital media artifacts, contributing to an expanded illusion of life. Higher degree of liveness implies more embodied experience of life. To illustrate the idea, I list several examples of animated artifacts in order of their respective degrees of liveness.

Playback Animation

As I have mentioned, all direct playback of the traditional film form of animation, including stop-motion animation, hand-drawn animation, abstract films, and CGI animation, is not live. The status is regardless of whether the presentation is a film or video projection, on a television screen, or in an Internet browser, as long as the whole presented content is playback of pre-recorded or pre-rendered materials stored in certain medium, such as a film reel, a videotape, a disc, a hard disk, or a memory board. It follows that a playback of the Pixar's animated short film *Luxo Jr.* (1986) on YouTube or the opening movie of an interactive multimedia websites is not live.

Slightly Live Animation

On the other hand, a playback of movie clips with very minimal interaction (just a little more than hitting the “play” button) on a website can be slightly live. A good example can be seen in the web advertising viral campaign at the website Comeclean.com. The multimedia website first presents its visitor with a subjective view of looking down at a sink basin and prompts a confessed message. Once the visitor types in a message, the website shows that the message is handwritten on the photographic palm of a hand and then is washed away with a cleansing foam in a real-time composite moving image. The display seems like a playback of live-action movie segments, which in fact is not a pure direct playback, but rather real-time composite imagery in response to visitor keyboard input (the confessed message typed by the visitor is superimposed on the palm in the segment in real time). Therefore, one should not presume that only computer-

generated visual content could be “live.” The liveness of prepared visual content, no matter live-action or animation, can be increased by real-time user intervention.



Figure 4.19 Screenshots of Comeclean.com where confessed messages input by users can be washed away

Fairly Live Animation

Meanwhile, a few interactive animated comics on the Internet or other digital media are fairly “live.” Examples include Han Hoogerbrugge’s works such as *Modern Living* published in Hoogerbrugge.com and Erik Loyer’s “opertoon” (a cartoon story one can play like a musical instrument) *Ruben & Lullaby*. In these animated comics, most visual content is direct playback of hand-drawn animated clips with simple variation in order of appearance according to user intervention. They involve very limited real-time manipulation on stored materials, such as color manipulation, image processing filter effects, timing or speed control, and the like. Yet, they all demonstrate possibilities of making heavyweight stored components fairly live by designing appropriate real-time user-driven manipulation.



Figure 4.20 A screenshot of *Modern Living* in Hoogerbrugge.com

Highly Live Animation

Most highly “live” animation phenomena can be found in the GUI of many computer systems. Exemplars are those commonly seen generic-styled screensavers, which can be traced back to John Whitney’s renowned abstract computer animation (Figure 4.21). There can be very limited or even no stored materials in the running of those screensavers. They are purely live. Meanwhile, animated GUI mechanisms in the Macintosh OS X system, such as the magnification effect in the dock, the genie effect of windows, bouncing of icons, and others, can also be called “live” animation because playback of stored content is only limited to still graphic imagery, like icons or fonts.

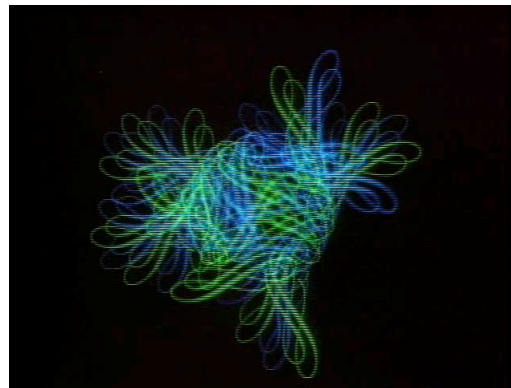


Figure 4.21 John Whitney’s *Catalogue* (1961)

Other alternative, but equally salient, examples of live animation can be seen in many works of digital art. One of the important installations in this area is Camille Utterback and Romy Achituv’s *Text Rain*. The work shows a projection of animated

letters falling like raindrops. A participant standing in front of the projection can move her or his mirror image shown on the screen to catch, lift, and then let fall any letters. During this embodied interaction, the only stored content is those falling typographic alphabets. The motion and reaction of letters, together with the participant's image, are all produced and composed in the presence of the participant.

Another interesting work worth mentioning is the aforementioned interactive outdoor advertising event *Interactive Wall: BIG SHADOW*. The event took place in the city center of Tokyo outside a commercial building, featuring a huge shadow play on the façade. The projection showed silhouette images of on-location players and a gigantic dragon in direct interaction, with the latter controlled by online web visitors in real time. The event successfully put people in different physical locations virtually together in terms of images, in which the dragon's silhouette is partly stored computer-generated content, but the people's silhouettes are certainly captured and transmitted on the location. The overall effect is an excellent demonstration of live puppeteerly animation. To sum up, the following figure illustrate the continuum of liveness labeled with the above exemplary works:

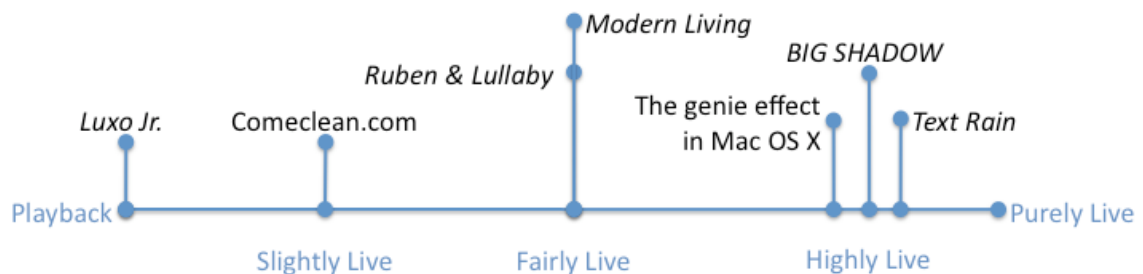


Figure 4.22 The continuum of liveness

Conclusion

This chapter has delineated the four principles of interactive generative animation, each of which garners the relevant knowledge and perspective of an indispensable part of the new animation phenomena, namely the observer, the image, the body, and the performer. The observer focuses on different types of liveliness presented in the phenomena, which blur the boundary between organism and machine. The image provokes material-based imagination at different levels, centralizing animated images in the nexus of perception, cognition, and bodily interaction. The body of the viewer or user is engaged in an array of interaction mechanisms with the artifact, embodying intention in motor action upon the environment. The viewer or user acts as the performer improvising with the artifact in various degrees of liveness and experiencing the contingency of life. All these defining characteristics are supported by the generativity and interactivity of the artifact. As mentioned in Chapter 2, the essences of the two notions include autonomous system and reactive instance, which engender different types of liveliness, invoke imagination, engage users in motor-sensory loops, and provide them with divergent and contingent experiences, by means of autonomous, reactive, metamorphic, and contingent kinds of animation phenomena intended by an author.

The text of the section “The Image” of this chapter, in part, contains material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2009a, 2009b). The dissertation author was a co-author of this paper.

CHAPTER 5

METHOD OF ANALYSES:

THE VARIABLES AND THE TAXONOMY

In the last chapter, I delineated the four principles defining generative and interactive animation. Each principle emerges from one of the four essential constituents of the new animation paradigm, representing knowledge and perspectives from the observer, the image, the body, and the performer respectively. To the observer, today's autonomous multimedia artifacts exhibit a holistic illusion of life, consisting of both prominent behaviors and emergent patterns, forming all-round liveliness. The reactive animated image mobilizes a motor-sensory feedback loop in the viewer, enabling an embodied understanding of meaning and sensation, provoking material-based imagination in terms of conceptual blends with elastic anchors. Meanwhile, the viewer also embodies meaning through bodily engagement with the artifact, comprising motion-based motor action taken upon an enduringly evolving environment. Finally, the viewer and the artifact co-participate in a live performance, reflecting the generativity, interactivity, and most inspiringly contingency of the system. All in all, a media artifact augmented with generative and interactive animation provides us with embodied experience of an expanded illusion of life, including autonomy, reactivity, metamorphosis, and contingency.

The Variables

In order to better encapsulate the notions of the four principles, I propose a set of qualitative variables describing the defining characteristics of a media artifact in the new animation paradigm, which has been briefly introduced in the previous chapter. The set includes the types of liveliness perceived from the observer's point of view, the levels of

imaginative understanding provoked by the image, the mechanism of bodily engagement with the artifact, and the degree of liveness of the co-performance between the user and the artifact.

The Observer: Type of Liveliness

With generative and interactive animation, a digital media artifact exhibits various types of liveliness, spanning from more prominent intentional or interactional behaviors, which concentrate the observer's attention on intention and dynamics, to relatively emergent patterns, which dilute one's attention over the ambience of a life-like environment. A lively media artifact should be able to show phenomena not only heightening a particular type of liveliness, but also encompassing an array of different behaviors and patterns. For instance, a lively cartoon film should have a figure prominently moving to show its own internal impulse and to react to external forces, as well as background elements subtly moving in rhythmic or stochastic patterns to complement the figure's action.

The Image: Level of Understanding

An animated image allows its viewers to understand the meaning or sensation it carries by conceptual blending. At the immediate level of understanding, the image acts as an elastic anchor, holding a viewer's sensory perception and motor action, and blends with the viewer's past sensorimotor experiences, to yield a largely automatic and unnoticed concept bridging the virtual environment and physical reality. At the metaphorical level, the image further provokes the viewer an imaginative story with a simple structure, usually alongside a well-known frame, resulting in more evocative messages. An animated artifact or system can be said as an embodied one if it incorporates tight immediate compression of cross-space mapping. Meanwhile, tight

metaphorical mapping implies an evocative and expressive artifact. In this regard, the water-level interface described in the previous chapter is an exemplar on both levels.

The Body: Mechanism of Engagement

Most animated multimedia artifacts invite viewer participation. A viewer takes bodily action upon an artifact and embodies one's intention through bodily engagement. The embodied intention or meaning hinges on first, the qualities of the viewer's bodily movement in time and space (e.g., speed and orientation), and second, a constant variation of the environment. These two qualities define the mechanism of engagement, reflecting the embodied friendliness or intimacy of an artifact to its human users. The least intimate mechanism is halted engagement, which involves no motion-based user input or enduring transformation of the environment. The moderately intimate mechanism is decelerated engagement, which lacks variation of the environment. The most intimate mechanism is continuing engagement, which entails both motion-based input from the viewer and enduring changes in the environment. Therefore, an animated interactive artifact with continuing engagement can be seen as a more human-friendly interface.

The Performer: Degree of Liveness

Generative and interactive animation is open to user intervention, like an improvised live performance. The degree of liveness measures this openness, and so the contingent nature of an animated artifact or how well a user can experience the contingency of life in the artifact. A high degree of liveness corresponds to what I call "highly live animation," which usually comprises very minimum amount of stored material content in playback. In contrast, low-degree liveness can be found in the type of "direct playback animation," which is always a direct playback of pre-rendered or pre-

recorded materials. Those in the middle of the continuum can be called as “fairly live animation.”

The above four qualitative variables and their possible values are summarized in the following tables:

Table 5.1 The variables and possible values

Type of liveliness	Prominent behavior	Intentional
		Interactional
	Emergent pattern	Rhythmic
		Stochastic
Level of understanding	Immediate	
	Metaphorical	
Mechanism of engagement	Halted engagement	
	Decelerated engagement	
	Continuing engagement	
Degree of liveness	Direct playback	
	Fairly live	
	Highly live	

With different possible values, these four variables form a multidimensional continuum for describing the features of any particular animated artifact with respect to those four essential facets of interactive generative animation. For instance, a video game, such as *Pong* (1972), might show an interactional type of liveliness, enable only immediate level of understanding, implement decelerated engagement, and present highly live co-performance. Meanwhile, a work of digital art, like *Text Rain* (1999), might show both interactional and emergent type of liveliness, provoke metaphorical level of

understanding, entail continuing engagement, and also demonstrate highly live improvisation.

These four qualitative variables, each of which is devised from one of the four principles of the new animation paradigm, constitute a descriptive framework for analyses of generative and interactive animation.

The Taxonomy

Toward analyses of generative and interactive animation, I propose an unconventional taxonomy, in which a thoughtfully composed corpus of media artifacts will be categorized according to an array of the four variables. These variables altogether span a multidimensional space, in which each dimension refers to one variable and the corresponding possible values distribute along the axis. In this abstract space, each artifact in the corpus articulated by this taxonomy will be tied to one specific coordinate point. In other words, each artifact will be represented by a profile of these variables in particular values, for example, (Type of liveliness: Interactional; Level of understanding: Immediate; Mechanism of engagement: Decelerated; Degree of liveness: Highly live), or (Type of liveliness: Interactional and Emergent; Level of understanding: Metaphorical; Mechanism of engagement: Continuing; Degree of liveness: Fairly live). These articulations render the proposed taxonomy radically different from a conventional classification with simple distinctive boundaries between its classes.

While the word “taxonomy” is usually concerned with classification of organisms, the use of the term here is relatively metaphorical. It identifies different species of animated artifacts, especially those mediated with machines or computers, which persistently evolve with our technologies, cultures, communities, and economies. Today, they have been developing into many different sorts of our everyday objects. We are used to regarding them separately as digital videos, video games, user interfaces,

digital artworks, multimedia websites, and the like. This folk classification can also be seen as the basic-level categories in digital media.

Based on studies in cognitive psychology and other fields, basic-level categories refer to the level of categorization at which children first learn about categories and objects (Lakoff, 1987, p. 14). It is the “level of distinctive actions” (Lakoff, 1987, p. 32). For example, people “play” video games, “browse” or “visit” multimedia websites, and “read” or “appreciate” digital artworks. It is also the level at which things are first named and used most frequently (Lakoff, 1987, p. 32), such as *video game*, *multimedia website*, and *digital artwork*. Further down the hierarchy subordinate categories, like *action game*, *role-playing game*, *simulation game*, and others come under *video game*. Conversely, working up results in certain superordinate categories, like *digital media artifact*. However, categorizing animated artifacts according to the basic-level categories, such as *video game*, *multimedia website*, or *digital artwork*, would definitely overlook many of their intricate qualities reflecting the long-standing human pursuits for the illusion of life.

The situation is like the classification problem of games raised by Ludwig Wittgenstein (Lakoff, 1987, p. 16). The classical theory of classification supposes clear separation between categories. Each category has a set of common properties shared by all its members. This classical model does not apply to *game* because we cannot find any common properties shared by all games. Yet, games are like families. A game might have something in common with another game, which might in turn connect to the third game by some other shared properties. These connections spread and form an intricate network among all games. Wittgenstein calls it “family resemblance.” By the same token, the new animation paradigm is like games and families. We might not find a single collection of properties shared by all members of *interactive generative animation* in the digital age. The variables I suggested do not mean the rules all animated artifacts must follow, but instead sketch out the directions and possibilities that members of this emergent kind are inclined to. Hence, the taxonomy based on these variables is not to subdivide the corpus

of animated artifacts into some mutually exclusive and pre-defined subcategories, but to disentangle and uncover subtle qualitative links in the corpus constituting the illusion of life.

Unlike a classical “box” classification (Bowker & Star, 1999, p. 10), the taxonomy I proposed does not precede the corpus itself. Instead, it is a vehicle toward convergence of qualitative analysis and humanistic interpretation of the corpus. The upshot is three-fold:

First, the variables establish us a novel but accessible vocabulary. Based on the theoretical underpinning described in previous chapters, the variables provide us with a set of shared analytical tools for describing and discussing the features of any artifacts in the emergent animation paradigm. Each of the variables marks a particular inclination of an artifact. On the one hand, this coding allows us to focus on the disposition of one particular attribute among a group of works. For instance, one can compare the emergent liveliness of a John Whitney’s analog computer animation with that of an ordinary screensaver on a desktop computer. On the other hand, also more importantly, the coding helps differentiate the nuances between convoluted attributes of an intriguing work, which may be overshadowed by the tight integration of form, content, function, and context of today’s multimedia artifacts. Consider the aforementioned water-level interface. It is primarily a utility application: a battery meter. Yet, its function is manifested in the form of computer animation and implemented in the context of everyday use of consumer electronics, altogether bringing out a provocative message. It seems that the function, form, and context are all tied together with the content. The variables help us extricate the knot by rethinking from four different perspectives one by one: how we perceive the lively movement, how we interpret the image, how we act upon the artifact, and how we cooperate with it. Through determining the value for each variable, we iteratively look into the connection between human perception, bodily action, and cognition.

Second, the taxonomy and the corpus are continuously evolving to reflect the emerging animation paradigm. With different values taken by the variables for each artifact, the corpus of artifacts distributes in the multidimensional space and sketches out an abstract terrain. By examining possible links between attributes and works, we look for patterns, norms, and peculiarities. Any unbalance in distribution would suggest a review of the corpus selection and composition. For instance, if the majority of the samples incline to the emergent type of liveliness over the prominent one, a researcher may consider whether it is a latest tendency of the animation paradigm, or only due to sample bias as some typical or salient candidates are excluded. That means the collection of samples is a dynamic and explorative process, aiming at a comprehensive and illustrative distribution of the properties of artifacts best demonstrating the new animation phenomena. Overall, the corpus represents a paradigm of generative and interactive animation, also a perspective on animated artifacts in the digital age. I admit that the corpus presented in this dissertation to certain extent reflects the vantage point and cultural baggage of the author too. I am open to any suggestions of artifacts to the collection with a primary objective of upholding the pervasiveness of the new animation paradigm in the basic-level categories of digital media (More discussions on the corpus is included in the next chapter).

Third, the taxonomy and the corpus constitute a new definition of animation in the digital age. In the taxonomy, the variable values determine the position of an artifact in the multidimensional space. The variables describe the interactional properties of the samples rather than the inherent ones. The former are based on our interactions with artifacts provided with our bodies and cognitive apparatus (Lakoff, 1987, p. 51). The latter are common properties defining the boundaries in classical theory of classification. As a result, the comparison between the distribution of the samples in the multidimensional space and their cognitively established basic-level categories reveals useful information for researchers and animation scholars. For example, *video game*,

multimedia website, user interface, digital artwork, and so forth, are digital media basic-level categories, because they are the most commonly used labels and each of them has its own cognitive image in our minds. However, they also largely overlap in the multidimensional space with only minor discrepancies, implying that they enable comparable perceptions and interactions. Hence, it is sensible to speculate that they might belong to a greater superordinate category: generative and interactive animation. I believe the results would shed new lights on basic-level effects (Lakoff, 1987, pp. 46-47) in digital media studies and suggest an embodied cognition approach to defining animation in the digital age. It is like understanding the concept of furniture, which exists at the superordinate level of categorization, requires mental images of basic-level objects, such as chair, table, bed, and the like (Lakoff, 1987, p. 52). Each type of furniture differs from others in terms of form, function, and motor requirement, but their interactional properties form “clusters” inclining to furnishing a living or working environment. Likewise, to define new animation in the digital age, we require a corpus of subordinate basic-level digital media objects orienting toward an expanded illusion of life.

In the next chapter, I shall talk about the composition of the corpus, followed by analyses of the artifacts with respect to the four variables.

CHAPTER 6

ANALYSES AND RESULTS

With an array of qualitative variables formulated in previous chapters, this chapter proceeds to give detailed analyses of a corpus of media artifacts. Each artifact will be described and articulated with a profile of values taken by each variable, formatted as a quadruple in order of (Type of liveliness; Level of understanding; Mechanism of engagement; Degree of liveness). Explication of the values with respect to the artifact then follows.

The corpus consists of diverse kinds of artifacts. The collection apparently has no specific focus on a particular type of media forms, yet there are reasons for its broad scope and particular selection. This intentional diversity serves the following purposes:

First, the corpus consists of works from a wide range of media categories, including films, video games, entertainment applications, user interfaces, digital art/performance, multimedia websites, and others. The variety of media types is intentionally included in the corpus in order to reflect the ubiquity of the new animation paradigm in today's digitally mediated environment. Every person in the modern world is, to a certain extent, exposed to these emerging animation phenomena.

Second, the corpus includes works from creators with very different motivations or ambitions. Some works might have practical purposes, such as user interfaces or multimedia websites. Some others aim to express poetic or critical messages, like in digital artworks or animated films. Some of them purely are intended as entertainment. Furthermore, a few of them were created for demonstration or experimentation of novel ideas. The vast differences in nature and objectives show that the emergent animation phenomena are independent of the genre, context, purpose, or meaning of a work. Works

of completely contrasting directions might be synchronous in projecting an illusion of life.

Third, there are both canonical and idiosyncratic works in the selection. Some of these works are typical examples of their own kinds (e.g., the GUI of Macintosh OS X, the entertainment application *Koi Pond*, and the multimedia website Comeclean.com) whereas some others are more eccentric, but still salient, examples (e.g., the interactive web comics on Hoogerbrugge.com, the digital artwork *Text Rain*, and the video game *Passage*). I include each of them in order to demonstrate that generative and interactive animation is not limited to either ordinary or extraordinary artifacts – it is a new animation paradigm in the digital age.

More importantly, the diversity in the corpus does not mean that the selection of samples is arbitrary. In fact, the following selection criteria persist in the process of analyses:

- The samples exemplify the new paradigm in terms of the properties captured by the variables.
- The samples are different from each other in at least one of the variables.
- Each sample is a representative from its own category only.
- The samples altogether span a widest terrain in the multidimensional space of the variables.
- The results would be somewhat different if any of the samples were not included.

All in all, artifacts from these basic-level media categories come together constituting the superordinate notion of generative and interactive animation.

The Corpus

Short Animation Films

Film⁷ is intrinsically a linear medium. No matter whether one is materially stored on plastic film, on videotape, as a DVD-video, or as a computer file in a hard disk, it presents its content in playback. It belongs to the category of playback animation. Although the playback of films affords very primitive interactivity, such as play, pause, fast-forward, reverse, and stop, nearly all animated films are considered as non-interactive in the narrow sense. They normally engage audiences in terms of visual perception but seldom incite bodily motor action, except very subtle body postures like reclining or upright viewing positions. Hence, the variable about bodily engagement might not be applicable to films. However, it is still informative to look at the types of liveliness shown in animated films, and the levels of understanding they provoke.

John Whitney's Arabesque (1975)

(Type of liveliness: Emergent; Level of understanding: Immediate and Metaphorical;
Mechanism of engagement: not applicable; Degree of liveness: Direct playback)

John Whitney, Sr. is well known for his innovative attempts in producing abstract computer animation composed of geometric patterns. His earlier works, like *Catalogue* (1964), were produced using his mechanical analog computer. After receiving an IBM research grant in 1966, he began to produce animation with the aid of digital computers (Russett & Starr, 1988, p. 180). His renowned works at this stage included *Permutations* (1967) and *Arabesque* (1975), which demonstrate his interest in harmony between computer vision and music. At a later time, he developed and experimented with an interactive audio-visual software system, which allowed one to compose visual and musical output in real time, producing so-called “visual music,” such as *Spirals* (1987) (Moritz, 1997; Russett & Starr, 1988, p. 26). As he put it, the system provides “a set of

⁷ The term “film” here refers to moving images in general, such as those in video, computer files, and plastic film.

controlled parameters that determine speed, color, shape, position,” and the like, which are also “connected with the pitch and duration of tone” (Russett & Starr, 1988, p. 26). In other words, the computer connects sensory perception, including the visual and the auditory, with real-time control. To me, Whitney’s interactive system is not only a manifestation of his aim of achieving ultimate harmony between visual and sound, but is also one of the earliest attempts to expand the notion of animation for the digital age.

Whitney’s *Arabesque* is an ideal example for analysis, because it stands in the middle of his creative investigation’s trajectory from analog computer animation to digital interactive animation, marking a turning point of his study in generative computer graphics from the traditional film form to interactive media. It serves as a good transitional representative of generative and interactive animation. Moreover, the title and the middle-eastern style music soundtrack by the Iranian-American santurist⁸ Manoochehr Sadeghi imply a strong association with certain culturally-derived aesthetic rather than purely formal geometric beauty, which would invite further imagination from most audiences, resulting in more levels of understanding. Meanwhile, the geometric visual form of the piece, just like other Whitney works, together with the synchronized musical rhythm, demonstrates a typical example of emergent patterns of liveliness like dancing.

Liveliness

The film starts with an opening credits sequence featuring a border of swirling motifs, which resembles arabesque – the outstanding form of Islamic art commonly found decorating the walls of mosques – in continuous motion. The fluid movement of the patterns throughout the film is emergent in the sense that one could not perceive any

⁸ Santur is a type of traditional Persian musical instruments.

prominent origin of impulse, center of gravity, or interplay between action and reaction. Instead, momentum seems to be distributed amongst visual elements over multiple different parts of the frame. All these animated components play a role in drawing the viewer's attention. Another characteristic of emergent liveliness the film presents is lack of noticeable progression and consequence in happenings. The movement in the film is phenomenologically and temporally reversible. That means reverse playback of the film would not result in significant change in the viewer's perception of liveliness. These features make the swirling curves a demonstration of emergent liveliness.

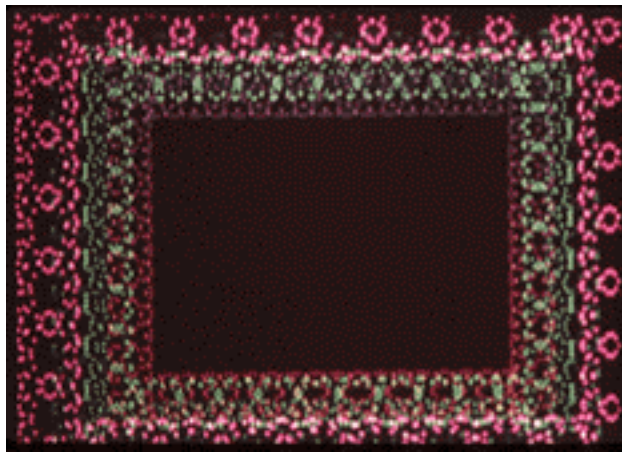


Figure 6.1 An opening screenshot of Whitney's *Arabesque*

Understanding

In addition, the swirling patterns are always in sync with the music soundtrack. The former moves increasingly vigorously in alignment with the latter. The complexity of the movement also increases with the number of soundtrack layers. To an audience, the film is a piece of choreographed work and the curves are performers. This understanding takes place immediately through a process of conceptual blending. At the very beginning of the opening credits, the curves move along the border rhythmically with the music, looking like a line of performers walking in sync on the stage. Later on, there are two sets of curves, like two groups of dancers, one at the up and down stage twisting slowly, and the other in the center moving quickly, each of which keeps pace

with the corresponding music track. When a viewer sees the patterns swirling on the screen and listens to the music, the synchronicity emerges from cross-space mappings with the observer's sensorimotor experiences of seeing, or even participating in, dance performances. (Although a viewer might only perceive a dance rather than actually taking part in the dance, the animation can still be regarded as a sensory-based elastic anchor if the action-perception coupling argument suggested by activation of mirror neurons is taken into account.) The outer-space relations include the analogy relation between patterns and dancers, between swirling lines and moving body parts, the category relation between soundtrack and music, and the analogy relation between screen and stage. They are selectively projected or compressed into the blend, in which patterns dance along with the soundtrack on the screen-stage. The following integration diagram illustrates the blend (Regarding the representations of elements and their relations in the diagram, please refer to "The Image" of Chapter 4).

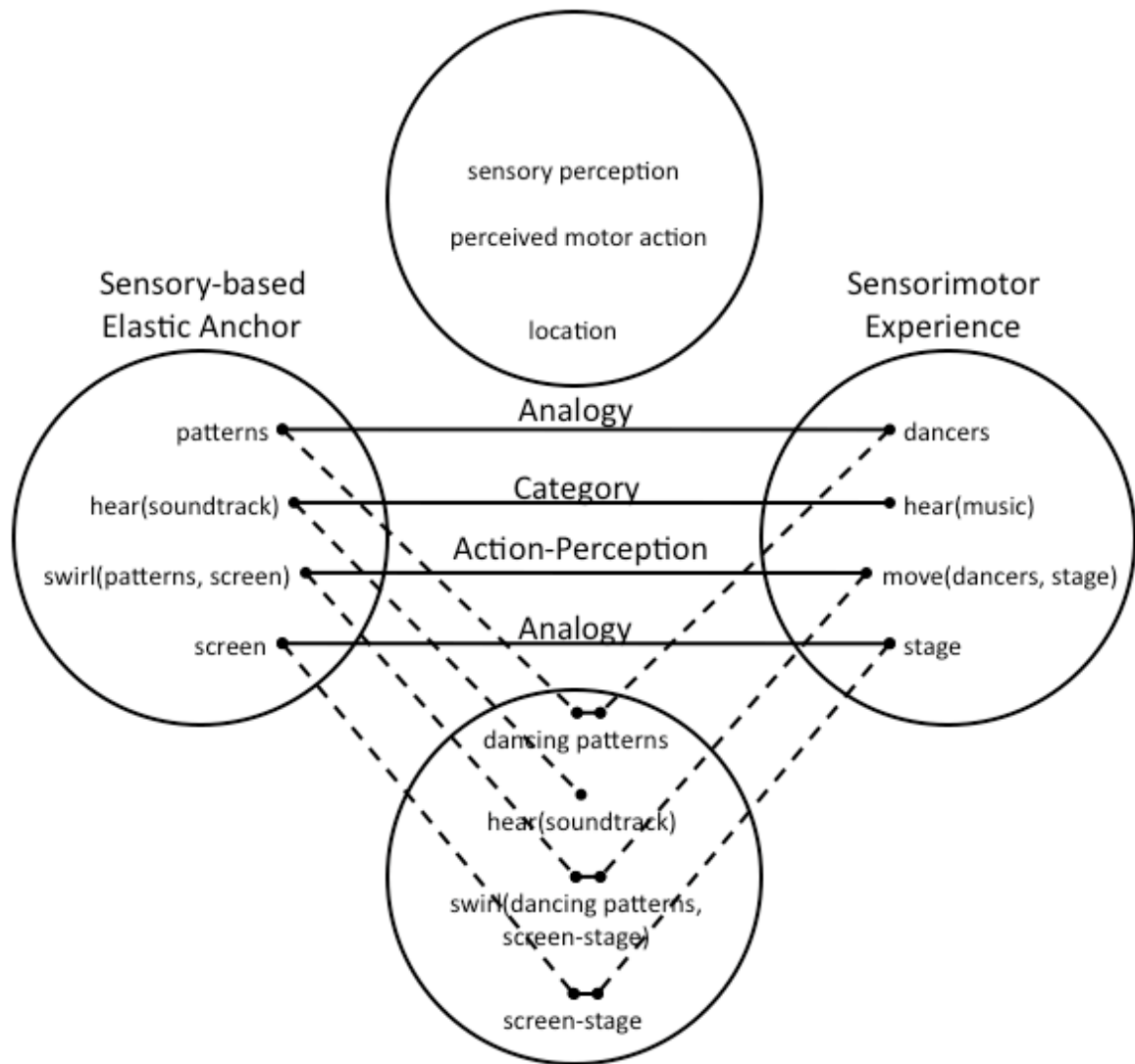


Figure 6.2 The immediate blend in *Arabesque*

Furthermore, the title “Arabesque” and the Persian instrumental (santur) music soundtrack remind audiences of association with Arabian context. They invoke the frame of Islamic art at large, including not only decorative patterns arabesque found in Islamic architecture but also Arabic dance (e.g., belly dance), as well as motifs of the related costumes. Given this frame, the patterns in the film resemble motifs or outlines of Arabic dance costumes in the viewer’s stereotypical impression of the Arabian, and the whirling movement is visually similar to the swing of bits and pieces of Arabic costumes in dance. Hence, the dancing patterns in the film provoke a blend with Arabian dance performance, outputting an imaginative Arabic dance in the dark with only vivid colored motifs of the

costumes being seen. This secondary blend, illustrated in the next diagram, provides audiences with a metaphorical, contextual perspective of abstract animation of pure geometric forms.

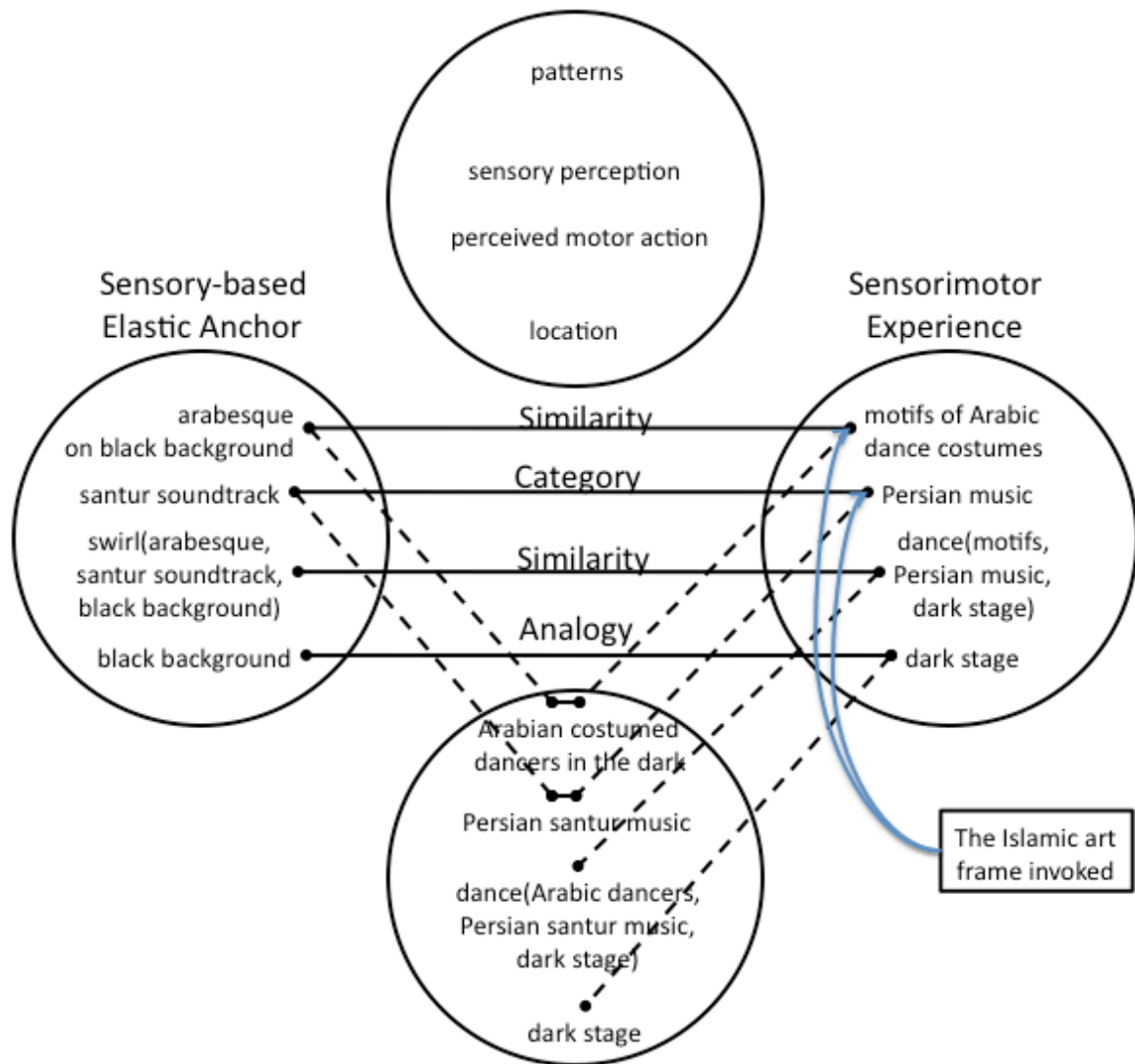


Figure 6.3 The metaphorical blend in *Arabesque*

Pixar's Luxo Jr. (1986)

(Type of liveliness: Intentional, Interactional, and Emergent; Level of understanding: Immediate; Mechanism of engagement: not applicable; Degree of liveness: Direct playback)

Pixar Animation Studios has become a renowned computer animation studio in latest two decades. The studio has produced over a dozen of massive feature-length animations, as well as numerous creatively explorative and technologically influential short films. *Luxo Jr.* is the first short film produced by Pixar and directed by John Lasseter, featuring a pair of anthropomorphic desktop lamps, one bigger and elder, the other smaller and much younger. The story starts with the lonely elder Luxo in the dark, woken up by a plastic ball, followed by disturbance from the fascinated and excited Luxo Jr., who later unwittingly and blunderingly bursts the ball. After feeling remorse for a short period of time, the baby lamp surprises his dad again by getting another even bigger ball. The simple but affective plot involving minimum number of characters (Luxo and his Jr.), certain suspense in the middle, and a little surprise at the end, not only makes the short a canon of its kind but also sets a typical story model for many following independent works of computer animation with characters. The film demonstrates a classical way of showing liveliness through anthropomorphic objects and constitutes some embodied understanding of body gestures.

Liveliness

As usual in character animation, the film features its characters in very prominent actions, imbuing originally still objects with both interactional and intentional liveliness. The opening focuses audience attention on a desktop lamp with a spotlight. A ball rolls in, coming to a halt beside the lamp. Everything seems to follow the laws of physical reality, until the lamp suddenly wakes up. The unanticipated movement surprises the audience and provokes them into taking the intentional stance, seeing the lamp having intention toward the ball. The lamp shows intentional liveliness. At a later time, when the ball bounces back twice, the audience is inclined to guess where the ball wants to go, until the baby lamp jumps in. While the dad keeps his eye on the kid chasing after the ball, and then looks downward on his blundering kid, interactional liveliness seems to

pervade. This kind of switching between the two prominent behaviors is commonplace in mime, as well as character animation.

On the other hand, emergent liveliness also persists throughout the short in terms of secondary animation, like the wave of the power cord following through the jumping of Luxo Jr. These movements do not captivate viewers' main attentions, nor do they contribute to any progression in narrative, but they provide a life-like ambience to the scene. They are indispensable for an all-round lively world. In making *Luxo Jr.*, the approximation of the wave motion by computational methods is believed to be one of the technological feats. With subsequent advancements in computer graphics after this work, computer animation films have become increasingly saturated with various types of emergent movements, like sway of grass straws, Tinker Bell-like dust effect, splash of water, and so on.

Understanding

As I have mentioned in Chapter 3, animation constitutes embodied understanding of sensation or intention, which is justified by conceptual blending theory and neuroscience accounts of mirror neurons. When an audience sees an anthropomorphic object performing, he or she immediately understands the intention or sensation by blending one's own sensorimotor experience with the perceived action. In the beginning of *Luxo Jr.*, the audience automatically follows that the lamp is "curiously" (a sensation) looking at (an intention) the ball. By simply mapping the light bulb with the eye, the cap with the head, one recognizes a common act of examining. The following figures illustrate this immediate blend.



Figure 6.4 A screenshot showing Luxo examining a ball

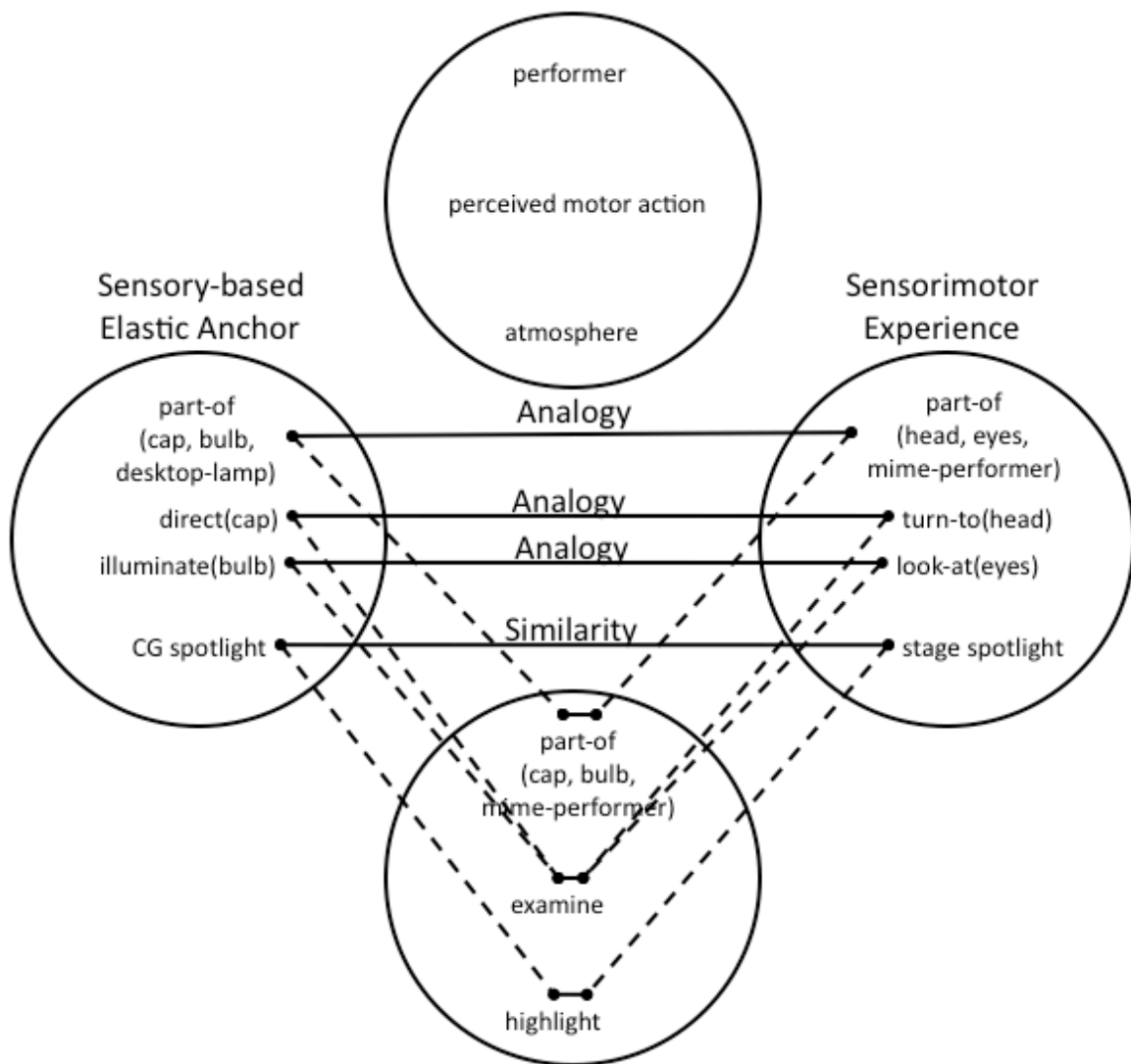


Figure 6.5. The immediate blend in *Luxo Jr.*

By the same token, when the lamp jumps to its feet, we know that it is struck. Some might say that we recognize the animation because it looks similar to what we saw before, for example, the reaction of a frightened cat. In fact, we understand that a cat is frightened all because we blend the perceived action with our own motor experience of shock. Since the blend is so frequent that just gone unnoticed, we immediately know there is a shock when something suddenly recedes. Moreover, the activation of mirror neurons that suggests the coupling of perceived and performed actions also provides promising reference to this immediate embodied understanding of sensation.

Arabesque and *Luxo Jr.* are both animation generated by the computer. The former, regarded as abstract film, exhibits emergent liveliness, while the latter, so-called character animation, shows more different types of liveliness. In contrast, the former provokes more levels of imaginative understanding than the latter. It may be because the non-representational form in *Arabesque* lends itself more readily to flexible metaphorical projection, whereas the photorealistic representation in *Luxo Jr.* confines viewer interpretation more to the director-intended meaning. The above analyses of these two salient works show that computer animation is not homogeneous in demonstrating liveliness and conveying meaning.

Video Games

Computer animation films, including Whitney's generative pure animation and many typical CGI films like *Luxo Jr.*, are only marginally a type of generative and interactive animation, because they are not real-time interactive in the sense of responding to user feedback. On the other hand, video games provide us with quite a few good examples of the new animation paradigm because they show many graphic elements in motion that are reactive and transforming in real time. The phenomenon is more obvious in subgenres of games such as action, sports, and simulation. Since there are innumerable examples in these groups, I select just a few because they illustrate the

variables clearly enough. They include one “classic” arcade game that is regarded as almost antique, but is actually still influential in terms of game mechanics, and one mini-game that looks like typical platform games, but is exceptional in meaning. Readers should note that the analyses here do not intend to focus on common parameters of game studies like rules and genres, nor do they attempt to revitalize the contention between the ludologist and narratologist camps. Instead, this section concentrates on how these games demonstrate the emerging animation phenomena in digital media.

Pong (1972)

(Type of liveliness: Intentional and Interactional; Level of understanding: Immediate; Mechanism of engagement: Decelerated; Degree of liveness: Highly live)

Pong was arguably the first arcade video game in the sports category, released by Atari Inc. in 1972. The game simulates a table tennis game with a rudimentary two-dimensional graphical representation of two paddles in opposite sides and a ball. The ball bounces back and forth between the two opposing paddles, which are controlled by two competing players or a player versus the computer, but constrained to move along only one axis to hit the ball. To me, the game is more like a table hockey game in that all movements take place in a two-dimensional plane rather than three-dimensionally as in table tennis. Although *Pong* is an early video game, its mechanics are very influential to many of its successors and variants. Today we could still find some mini games that have implemented similar gameplay, such as *Break' Em All* (2005) released on the Nintendo DS platform and *Touch Hockey* among many iPhone applications. I choose this game not just because of its legacy in the history of video games, but actually due to the fact that it does distinctively demonstrate the features of generative and interactive animation.

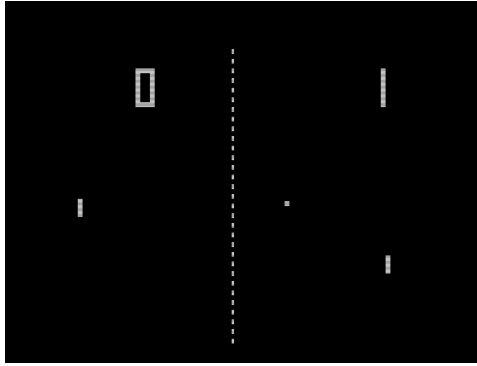


Figure 6.6 A screenshot of *Pong*

Liveliness

The classical game features three prominently active characters, the two paddles and the ball. They contribute to some quite prominent liveliness. First, the rivalry between the paddles in the game definitely shows an interactional type of liveliness. To an observer, both paddles often act like racing toward the elusive ball, which on the contrary seems always like getting through either one of them toward the bottom line. These motions make visible the dynamic tension between these characters. Second, the act of a paddle, even though it might be computer-controlled, sometimes would surprise and provoke the observer into taking the intentional stance and speculating its tactics. That means dynamic tension is not the unique force governing a paddle's behaviors. Sometimes a paddle would perform suspiciously and unexpectedly. For example, it might overshoot the actual touch point of the ball by a little and so quickly recede. It seems to make mistakes, even if this computer opponent might be actually programmed. Furthermore, in some later variants of the game, the computer contender could even suddenly change its direction when hitting the ball in order to play tricks like adding spin onto the ball. All these unpredictable acts reveal certain extent of intentional liveliness of the system.

Understanding

Like many other action games after it, immediate blends take place in *Pong*'s game mechanics, entailing the motor-sensory connection between player action and animated visual feedback. In the original arcade and home version of Pong, the player is provided with a rotary knob through which one is able to move the corresponding paddle along the bottom line to hit the ball represented by a white dot. When the player turns the knob, one could see the paddle on his or her side moving accordingly to bounce off the ball. How fast and how far the paddle goes instantaneously reflects how the player acts on the knob. This motor-sensory feedback loop makes the virtual paddle become an extension of the player's body, constituting an elastic anchor for an immediate blend with the player's sensorimotor experience of playing table hockey. In the blend, the identity relation between the paddle and the player is compressed into an imaginative whole in the virtual space (it is the reason why some players incline their bodies with the paddle when playing the game); the visual representation relation between the white dot and a small puck is compressed into a virtual puck; the analogy relation between the screen and a table is compressed into an imaginative slippery table; the player action and sensory perception are also selectively projected. The following diagram illustrates the structure of this blend.

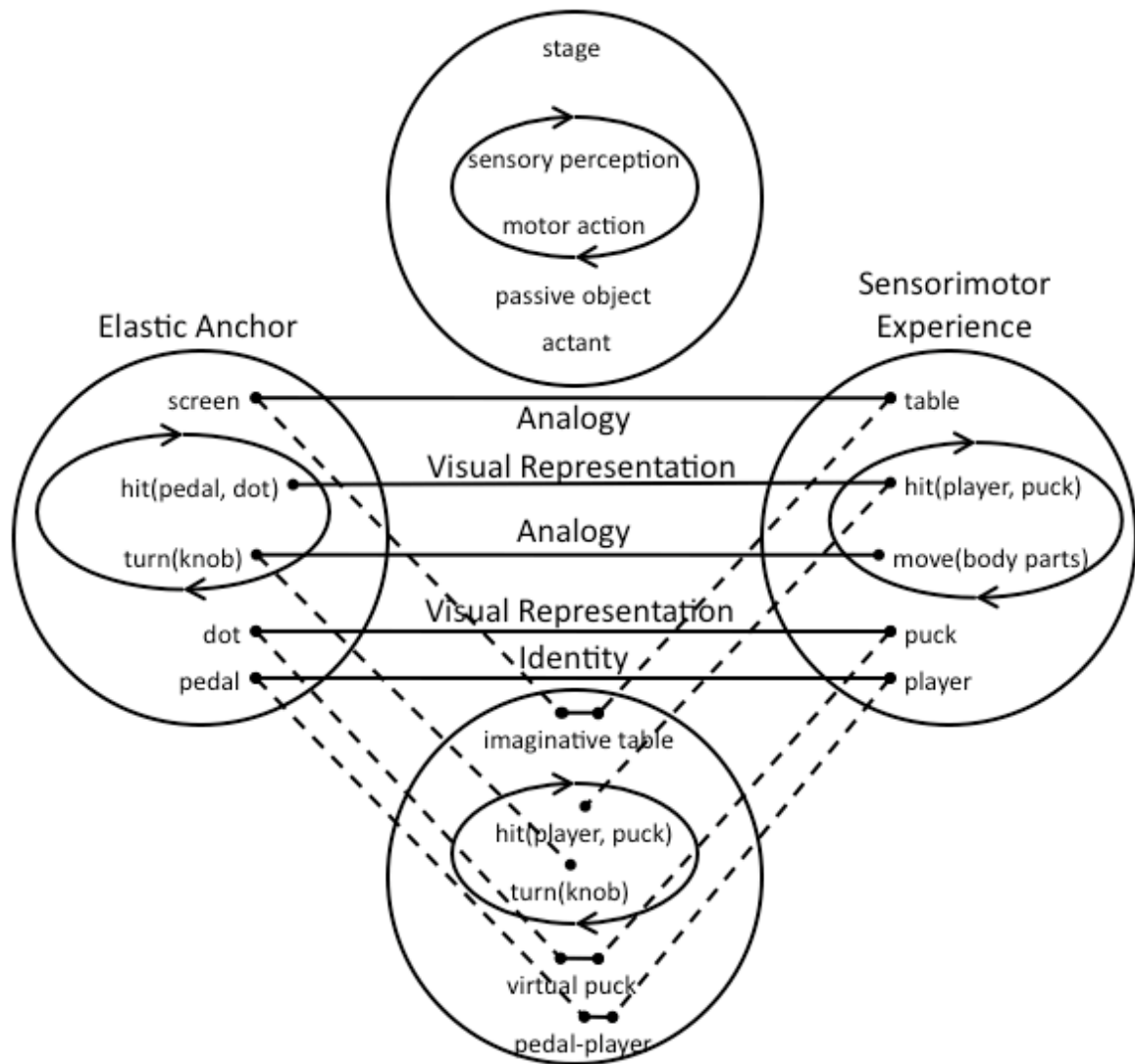


Figure 6.7 The immediate blend in *Pong*

Bodily Engagement

Since most action games incorporate animation that is real-time interactive, we could look into their mechanisms of bodily engagement. As an action game, *Pong* engages its players through not only sensory perception but also bodily interaction. The player's hand connects to the virtual paddle through the rotary knob, which enables the player to control the sliding speed of the paddle. The hand's turning act determines the paddle's moving speed. Moreover, in hitting the puck, the moving direction and speed of the paddle simulate adding spin and so affect the bouncing pattern. In other words,

motion qualities of the player's motor input have significant effect on the animated feedback. The input to *Pong* is motion-sensitive.

On the other hand, the gameplay of *Pong* follows the ordinary two-on-two ball game. This kind of game requires the tight alternation of the opposite sides taking turns to act. In case any one side withdraws from the game, it would come to a halt. In *Pong*, if the player stops taking action, the immediate result would be very likely that the player loses one point. If the player further refuses to kick off next point, the game just pauses as long as the player resumes the action. It is not sensible to say the environment would continue persistent transformation and give rise to divergent outcomes without the player. Hence, *Pong* falls short of facilitating an enduring interaction, and only affords decelerated engagement.

Liveness

Pong, as a sports video game, like its actual counterparts, can be seen as a form of performance. The animation phenomenon shown in this performance is also highly live, because the amount of stored visual content is kept to a minimum. The paddles and the puck are only represented by a bar and a dot, which need not be pre-rendered and stored, but just generated on the fly. The motion of the computer paddle and the puck is also simple enough that can be computed in real time. The high-degree of liveness in *Pong* implies that the system is very contingent, and the game can be regarded as an improvised performance between the system and the player. The player would enjoy being both the performer (when hitting the puck), and the audience (when seeing the opponent running for it), of the game show, especially for an experienced player.

Passage (2007)

(Type of liveliness: Intentional and Interactional; Level of understanding: Immediate and Metaphorical; Mechanism of engagement: Continuing; Degree of liveness: Fairly live)

The disposition of *Pong* in terms of the variables reflects the status of most action or sports games. The action of opponents or antagonists commonly shows interactional and intentional liveliness, because they are programmed to defeat or disturb the player, such as the ghosts in *Pac-Man* (1980). The control allows the player to take action directly in real time, enabling immediate embodied understanding, like the Nintendo Wii Sports series. Once the player stops action, the game would not advance to next levels. The animated graphics is usually highly live. Meanwhile, other genres of video games might show certain different patterns. For example, sandbox games like *Electroplankton* for the Nintendo DS device do not involve competition, and the intentions of animated components might not be clear. Simulation games such as the *SimCity* series would show enduring and diverging transformation of the environments even without any player input for a certain period of time. Some serious games or expressive games would further provoke metaphorical levels of understanding, like Jason Rohrer's mini game *Passage*.

Jason Rohrer is an independent game designer, who has made several mini games that have been very well received on the Internet and in both the indie game and academic game studies communities. One of his most notable works is *Passage*, which is an expressive game written for meditation on life and death and as a reminder of mortality. As Rohrer puts it in his creator's statement, "*Passage* is a game in which you die only once, at the very end, and you are powerless to stave off this inevitable loss" (Rohrer, 2007). He believes that computer codes could "make us cry and feel and love" and intends to turn video games into artistic vehicles for exploring the meaning of life (Fagone, 2008). Hence, *Passage* is imbued with evocative metaphors of "life is a journey," presented literally as a long screen of a maze. A player starts the game with a character walking alone in a maze with many obstacles. The character sprite's fixed position on the screen shifts from the left edge to the right gradually within the five-minute game time. Simultaneously, the sprite is depicted as increasingly older (going bald and grey, becoming hunched, etc). Along the way, the player accumulates points.

One might collect treasure chests for additional points, or walk with a companion for double points. However, once the character bumps into his spouse, the couple walks hand in hand and becomes less agile to collect treasures. The player has to make a trade-off. Meanwhile, the most poignant part is that no matter how many points one got, the character has to die after reaching the right end in five minutes.



Figure 6.8 A screenshot of *Passage* with the player character at the beginning of his life journey

Liveliness

Unlike most other maze games, *Passage* does not involve any lively adversaries who disturbingly challenge the player character. The only character other than the player character is his spouse, who contributes to the minimum interactional liveliness of the game. At the beginning, the lady just stands still in the maze. Once the player walks close enough to her, she snaps to the player and they start walking side-by-side. The couple shows a magnetic type of action and reaction in the game. Besides, when the player character, with or without his spouse, is exploring down south into the maze, the behaviors would elicit the intentional stance from an observer, speculating whether there are treasures nearby. In other words, the resulting animation also shows intentional liveliness. Unfortunately, the work did not demonstrate any emergent patterns.

Understanding

Although *Passage* only involves two characters and a passage, there is a wealth of metaphors provoking multiple levels of understanding. Apparently the work looks like a primitive maze game with the player character wandering and searching around. Yet the

presentation of *Passage* differs from an ordinary maze game in at least two folds. First, the screen is intentionally made to be exceedingly wide such that only a horizontal slice of the maze can be seen. Second, at the early part of the game, the region at the far right end of the screen seems to shrink and blur. Those visions would become clearer and clearer only when the character moves forward. These nuances render an immediate blend in *Passage* much tighter than in most maze games like *Pac-Man*, or even platform games. In *Passage*, when a player presses the arrow keys on the keyboard, the screen scrolls and reveals more walkway ahead of the character, and also the player. Compared with a fixed overview of the maze in some maze games, this incremental unfolding is more reminiscent of our sensorimotor experience of exploring an unfamiliar walkway in which the further you move, the more you see. This motor-sensory connection mobilized by the animation acts as an elastic anchor to the immediate blend. The identity relation between the player and the character is compressed into an avatar concept, an embodied extension of the player in the virtual space. The representation relation between pixel graphics of the maze and an unfamiliar walkway is compressed into a virtual passage. The output is a new concept of exploring an imaginative maze, as illustrated in the following diagram.

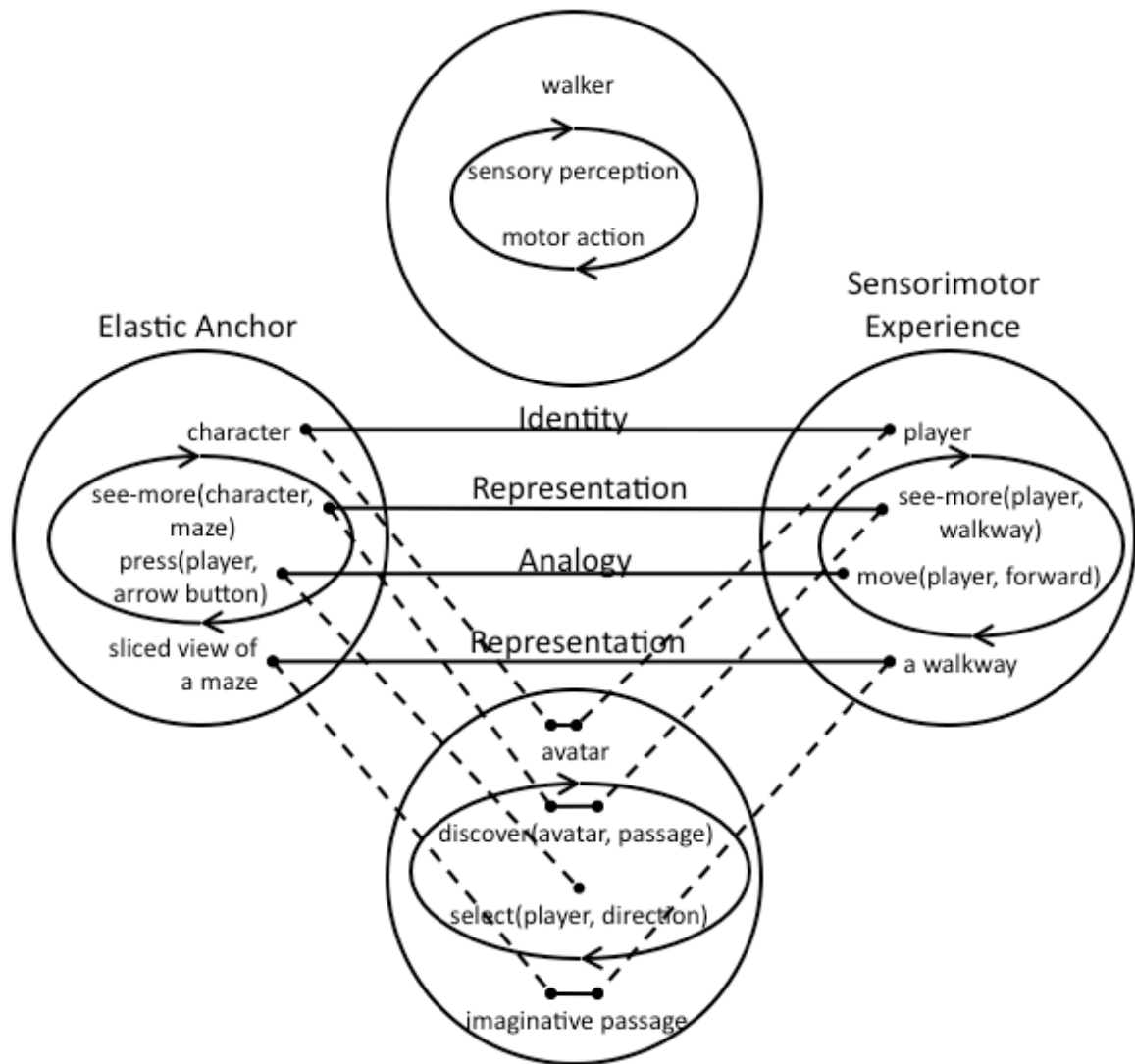


Figure 6.9 The immediate blend in *Passage*

As mentioned, *Passage* is imbued with metaphors related to the basic “life is a journey” metaphor. First, the game presents a long horizontal slice, and the character gets older and older as time goes by. When the character walks from left to right in the passage, the prevalent metaphor of space as time projects the journey on the wide screen to the lifespan of the character. At the end, the character approaches the right edge and then leaves only a gravestone, a conventional symbol of death. The five-minute screen journey represents the character’s short life. Second, in the course of his journey/life, he will see a lady. One could approach her making a life-long companion, or just leave her alone. The two options correspond to two fundamentally contrasting kinds of living. In

the former the couple becomes unable to explore the narrow alleys down south and collect rewards that seen nearby. They could just enjoy the mainstream passage. This limitation echoes the common frame of marriage, bringing out the author's intended imaginative perspective on how marriage limits possibility. These rules and constraints form the secondary blend of the game with a story of life and death. As shown in the following diagram, the analogy relation between the screen journey and a life with marriage, and that between a narrow sideway and a slim opportunity, are compressed into the new blend, together with the representation of death by gravestone. In the blend, the message: "till death do us part!" is invoked. While this diagram only refers to one possible imaginative blend, the work also provokes other understandings such as the adventurous and treasure-rewarding bachelor life. The analysis here only emphasizes one primary interpretation of the game.

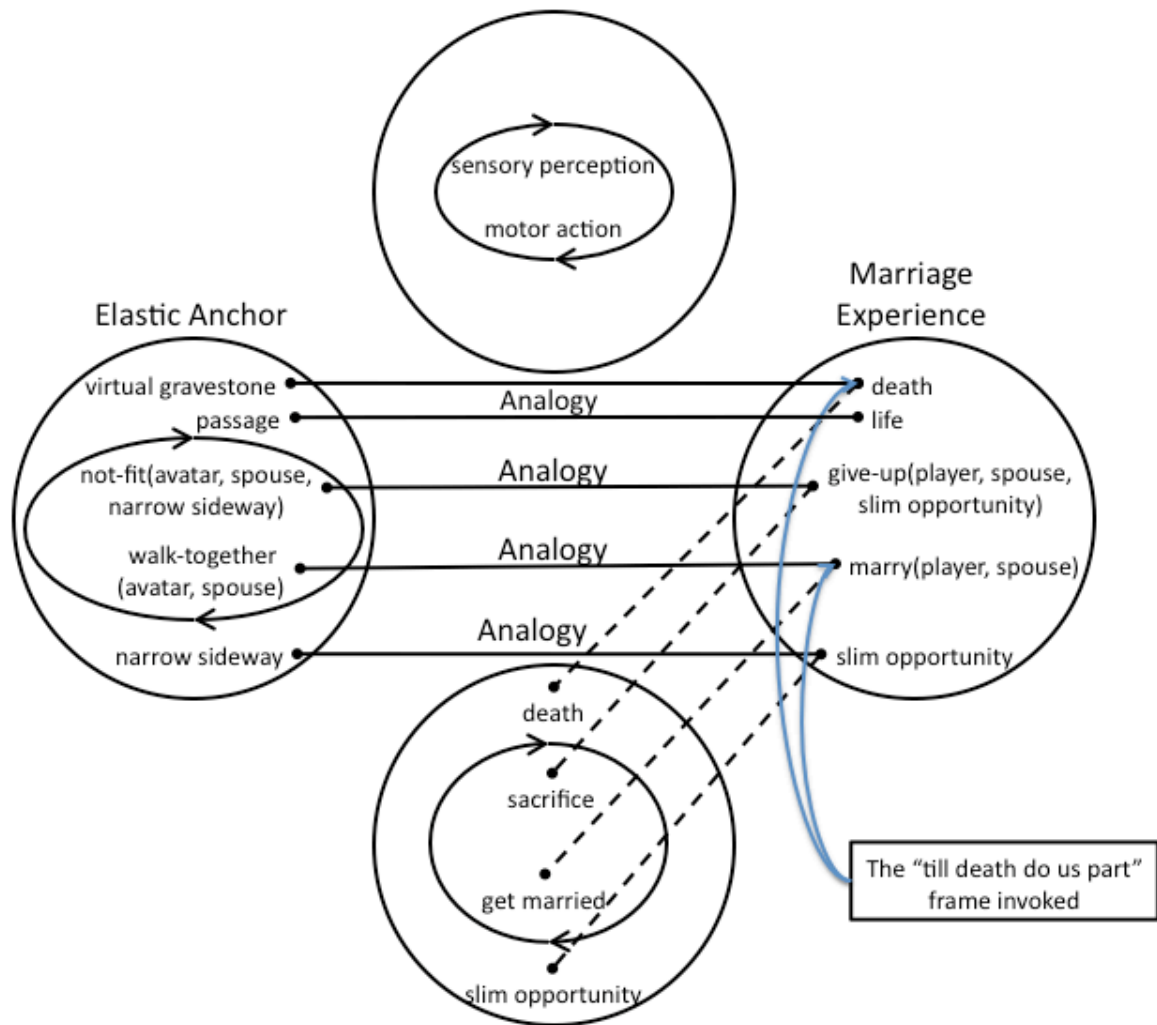


Figure 6.10 A metaphorical blend in *Passage*

Bodily Engagement

As a mini maze game like *Pac-Man*, *Passage* allows its players to control the motion of the character in only four directions, namely up, down, left, and right. In its computer version, players use arrow keys to move the character. One might say this input is not very motion-sensitive because qualities of finger motion like speed or direction do not have significant effect on the resulting animation, except maybe on the change of moving direction. In the iPhone version, however, the arrow keys are turned into four touch-screen buttons allowing a player's finger to just "slide" on the screen to control the direction. This button arrangement makes the character sensitive to finger motion

qualities, including motion path and speed. The player feels like orchestrating the character in motion by moving finger. This fairly motion-based input can also be achieved in the computer version if the touchpad (a basic input device of laptop computers in today's common configuration) is taken into use. Hence, the motor input to *Passage* is fairly motion-sensitive.

While some mini maze games might have motion-based input, the mechanisms of bodily engagement are mostly decelerated because the game environment would not advance once the player drops the control. This is not the case in *Passage*. If the character stands still, his position on the screen would shift toward the right at regular time intervals. That means even though the character's absolute position in the maze does not change, the screen position changes due to camera movement. In the meantime, the rendering at the right side of the screen is getting clearer and clearer, while those at the left side is becoming hazy. Hence, the game environment still shows ceaseless transformation without any user action. Apart from a continually changing environment, the character also gradually transforms from a young look to an elderly no matter whether he moves or not. If the player takes over the control again, the camera keeps track of the character and the screen would display properly changed and transformed visuals accordingly. Therefore, *Passage* engages its players continually and divergently in transformation.

Liveness

Compared with rudimentary action games like *Pong* or *Pac-Man*, *Passage* has more stored visual assets (sprites, terrains, and audio as well). The pixel graphics of the characters and those quite a few treasure chests were made and stored in advance. The structure of the maze was defined first, while the rendering, especially the shrunk and hazy effect, is likely to be generated on each instantiation, because it depends on how a player navigates in the maze. Furthermore, the compositing of all these elements into the

screen display is definitely done in real time. In other words, the only stored visual materials are still images of some pixel graphics. There are contingencies in playing out the game, resulting in a fairly live improvisation between the player and the game. When the character is wandering, the player is a performer. When the player notices the changing environment and character appearance, one becomes an audience too.

There are a myriad of video games. They are commonly classified into different genres whose boundaries in fact are not clearly separated. It might be hasty to say that action or sports games should show interactional liveliness and involve motion-based input, simulation games would present enduring environments, serious or persuasive games might provoke more levels of understanding. As *Pong* and *Passage*, both of which can be seen as mini action games, behave quite differently in terms of the variables, looking at video games from the perspective of the new animation paradigm sheds us new light on game studies too.

Entertainment Applications

With the advent of hand-held devices like the Apple iPhone, there has been an explosion of mobile applications in the consumer market in recent years. Applications constantly come up from various genres, including games, entertainment, social networking, productivity, lifestyle, travel, news, and many others. The games category consists of numerous applications that are created less for ordinary life and more for amusement. These applications run like video games on the mobile device, which have been generally covered in the previous section. Applications in other categories have more practical or serious purposes related to our everyday lives. For example, social networking applications enhance one's social communication with others regardless of time and space, productivity applications allow users to continue their office works on mobile devices, lifestyle applications enrich one's quality of life, travel applications provide travelers with many useful information or convenience, and so on. Between

amusement and practicality lies an emergent kind of entertainment applications, which explore to provide extraordinary, digitally mediated experiences enabled by emergent interactive multimedia technologies. They do not pre-define any explicit goals or rules of play, nor do they assume any application domain. Some of them reproduce everyday embodied experiences, such as popping bubble wrap, squashing cockroaches, gazing at a fish tank, and the like, in digital media. A few others simply create pure novel experiences we never have, such as *Spawn* mentioned in Chapter 3. To me, many of these applications are actually experiments with creating illusions of different forms of life. This section talks about a typical and good example of such an application.

Koi Pond

(Type of liveliness: Interactional and Emergent; Level of understanding: Immediate; Mechanism of engagement: Moderately decelerated; Degree of liveness: Fairly live)

Koi Pond was one of the best selling iPhone entertainment applications of 2008, released by The Blimp Pilots. The application is a computer simulation of a pond with colorful swimming koi fish beautifully rendered. When a user runs the finger across the touch screen, the koi dart away and water ripples spread out. After a while of rest, the koi come back again in tranquility. Together with situated ambient sounds like breezes, rains, or frogs, the application aims to immerse the user in an illusive life-like virtual environment (please see Figure 4.3).

Liveliness

The motion of the koi is likely to be implemented by clones of Craig Reynolds's *Boids*. Each koi seems to swim freely without bumping one another. There is no one major prominent force steering the fish. Instead, each of them leads their own way. The overall balance of forces projects an emergent type of liveliness. Meanwhile, the application allows a user to set weather condition, such as breezes and rains, which would

disturb the water surface. The resulting water ripples come and go, here and there, now and then, without making irreversible consequences to the pond, adding extra emergent liveliness to the virtual environment. On the other hand, when the user disturbs the pond by touching, a repulsive impulse drives all the koi away. If the user holds his finger still on the screen for a while, the touch point would turn into a center of gravity attracting all the fleeing koi back. Yet, any sudden movement of the finger would make them dash away again. By shaking the device, food is dropped into the pond, and they would come back again. This interplay of attraction and repulsion noticeably forms interactional liveliness in the pond. The emergent pattern of motion and the prominent behavior of interaction complete a holistic illusion of life.

Understanding

Making use of the latest multi-touch technology, *Koi Pond*, like many other handheld applications, is able to constitute an immediate level of understanding. When a user touches the screen and runs fingers across it, he or she inevitably and automatically recalls the sensorimotor experience of gazing into a pond and dipping fingers into it. This analogy of bodily action, also a vital relation, is tightly compressed into an inner-space similarity relation, with a major clash of the tactile experience being resolved in an immediate blend. Meanwhile, the user sees the reaction of water and koi in terms of the graphics on the screen and the sounds from the device, which projects a life-like counterpart to an actual pond environment. This representation relation is also compressed into a conceptualization of virtual ponds. Elaboration of the blend results in a nearby, accessible, and ready-at-hand lively virtual world in which the koi, water, and lily pads are all reactive to our finger motion. The following diagram shows one major phenomenon of the koi reaction.

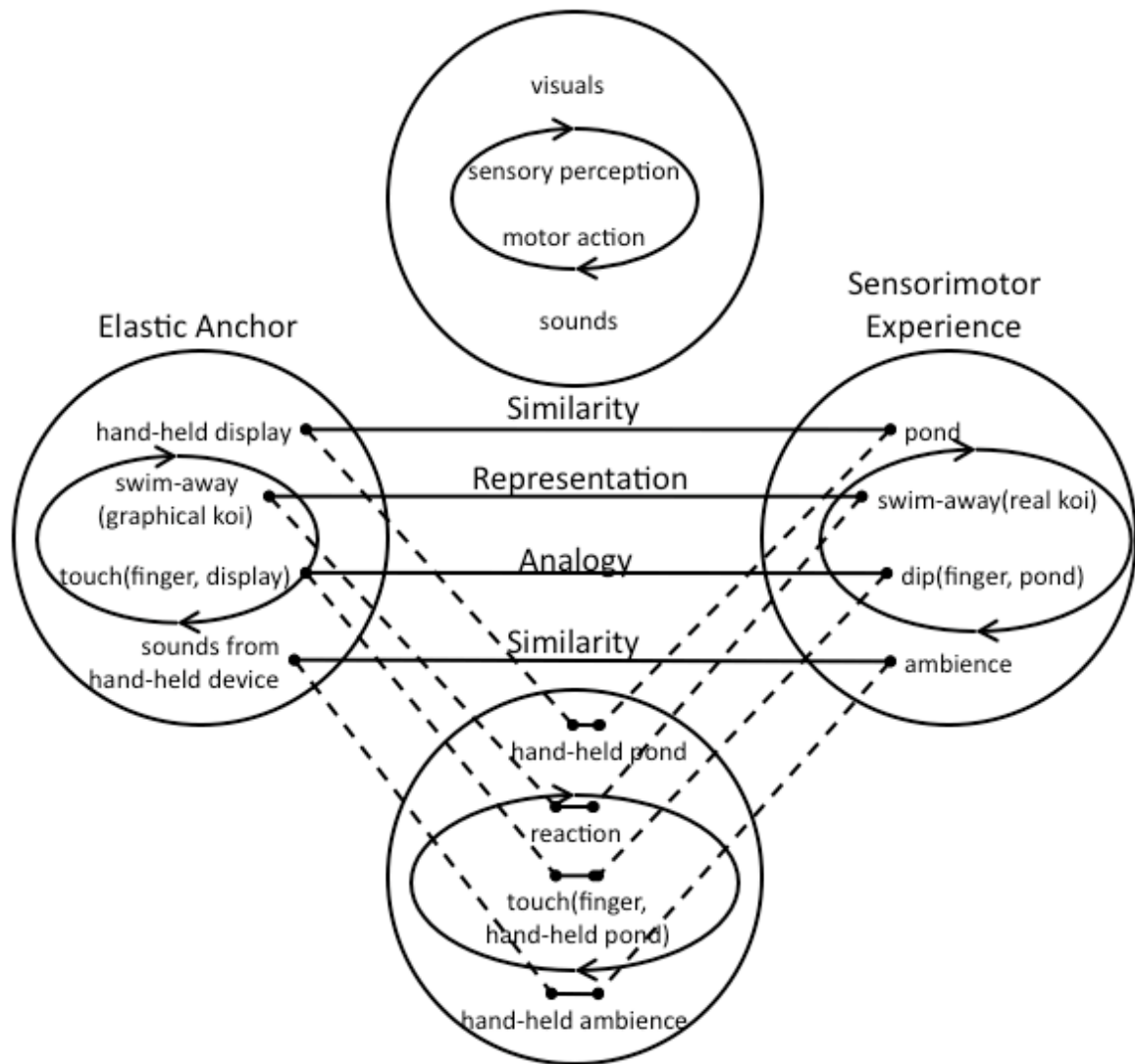


Figure 6.11 The immediate blend in *Koi Pond*

Bodily Engagement

The virtual world now featured in *Koi Pond* engages its users in a fairly “welcoming” environment. On launching, it invites one to touch and run fingers on the screen by showing beautifully rendered swimming koi. Various qualities of finger motion, including number and position of contact points, the pressure of touch, speed, motion path, frequency, and the like, determine what the next water imagery would look like, or which direction the koi would go. Hence, the motor input in *Koi Pond* is definitely motion-based.

Besides, the application does not simply wait for input and react, but instead shows its continuous “becoming” all the time. With koi constantly swimming around and water ripples in case of rains and breezes, an illusion of continuous variation is created. It seems to be an enduring environment. In fact, the pond makes no actual advancement in status. For example, the koi would not grow nor die, water keeps crystal clear even without cleansing, rains or breezes are turned on or off explicitly by the user only, and water would not overflow even raining ceaselessly. Hence, the case of *Koi Pond* is just marginally like a fully enduring environment, and the engagement is moderately decelerated and partly continuing. To make the application more self-evolving, one of the promising way, like *SnowDays* or *Sun Dial* mentioned in previous chapters, can be to reflect the time of day or the weather of the user’s geographical location on the virtual pond, with information live fed from the Internet. These methods can make the virtual pond really enduring.

Liveness

The real-time simulation system in *Koi Pond* enables its users to partly control the koi behaviors, like when and where the koi gather or disperse, and lets them gaze into the resulting animation. Users of *Koi Pond* are both orchestrators and audiences of the improvised performance. In the application, the stored components include the graphics of all the koi, the tiles, and the lily pads, while the koi’s deformation and motion, water ripples, changes of color tone, and so on should be generated in real time. Hence, the improvisation is fairly live, reflecting a certain extent of generativity, interactivity, and contingency of the application.

The variables’ values of *Koi Pond* position the work in the middle ground of the multidimensional space, where all four principles are moderately demonstrated. While the mobile application market is booming lately and is likely to flourish in future, this

domain provides us with a huge archive for studying the tendency of the new animation paradigm, which deserves continuation in future research.

User Interfaces

Films, video games, not to mention entertainment applications, are generally seen as forms of amusements. Yet, the new animation paradigm by no means refers only to amusement. In fact, one of the major application domains of generative and interactive animation is conventionally very productivity- and usability-oriented. That is interface design. Since the demonstration of Ivan Sutherland's *Sketchpad* (1963), Douglas Engelbart's early computer mouse, and others, user interfaces have become one of the indispensable constituents of computer systems, giving rise to the area of study HCI in the computing discipline. In addition, the idea of direct manipulation, together with the innovative designs from Xerox PARC and Apple, have steered user interfaces toward a graphical presentation. In recent years, the advances in technologies like computer graphics and interactive multimedia have even "animated" most of the user interfaces, not only on personal computers but also of many mobile devices and electronic gadgets. This section starts with discussing a very typical and well-known example of GUIs, the Macintosh OS X environment, followed by another relatively peculiar interface of a Japanese cell phone.

The GUI of Macintosh OS X

(Type of liveliness: Intentional and Emergent; Level of understanding: Immediate;
Mechanism of engagement: Largely halted; Degree of liveness: Fairly live)

In the early 1980s, Xerox, Apple, and others, respectively started to offer graphical user interfaces for their own computer systems. Just a couple of years later, Apple released the first version of her most successful series ever, Macintosh OS 1.0. The GUI of this initial version already incorporated most basic concepts of today's interface

design, including the desktop, folders, files, the trash, as well as the typical windows and icons. The similar framework has been adopted in many later variants and clones, with major front-end advances in number of colors and graphical details only, until the release of Macintosh OS X around the turn of the century. The brand new operating system imbued its interface with many animated visual effects, including the bouncing effect of icons, the “genie” effect of windows, and the magnification effect of the dock (alias icons grouped at the periphery of the screen). These effects demonstrate the characteristics of generative and interactive animation, creating an expanded illusion of life.

Liveliness

As described in Chapter 4, icons, windows, and the dock in the GUI of Macintosh OS X exhibit both prominent and emergent types of liveliness. The prominent behaviors are mainly performed by two lively widgets, namely bouncing icons and stretching windows. An icon of a launched application not in user focus sometimes would pop up and down in the dock, because it has a message or request for the user. This unanticipated but arresting action seems like someone jumping up and down in order to arouse others’ attention, inciting the user to take the intentional stance and to think about what happened to the application. Moreover, a minimized window, in case of a user click, would stretch and twist into place like a playful genie coming up from its oil lamp amazingly. On seeing this “body gesture,” a user likely wonders the reason behind this act. The window seems to “show off” its magical power that previous counterparts do not have. These erratic behaviors from these extraordinary widgets project GUI users an intentional type of liveliness.

Besides, the GUI also shows emergent patterns in the magnification effect of the dock. The dock contains a list of user-collected alias icons made easily accessible to the user. When one moves the pointer across the icons, they scale up and down successively forming like a wave traversing with the pointer along the dock. In other words, the user

feels like moving the wave in a stadium. Yet, this wave is always reversible in motion, and the pattern is emergent.

Understanding

Like what *Luxo Jr.* and some other character animation did on still objects, the GUI of Macintosh OS X anthropomorphizes its widgets. When a user moves across the dock, the icons scale up and down successively. If the user clicks on one, the icon would respond with bouncing restlessly. The user understands that the application is beginning the process of opening by immediate blending. The user effortlessly blends the action of running across icons, clicking, and seeing bouncing feedback, with socially recognized forms of body language for anticipation: restless shifting between upright and reclining positions, or even jumping up and down, as experienced in one's life competing for candidacy. The outcome is a new concept of calling a responsive and attentive application in.

Meanwhile, if the user clicks on a minimized window, it would “enact” a playful genie, stretching and twisting into place. The user knows the connotative meaning when blending the click and the perceived twist with the genie from film or television. In the blend, the motor-sensory loop of click and twist has a counterpart in our sensory experience of seeing a cartoon film (e.g., Disney's *Aladdin*), in which someone rubs the oil lamp and then the genie comes out. The analogy relation between the user click and the rub is compressed into an imaginative act of calling. The similarity relations of the window twist to the genie twist, as well as the desktop to the film background, are also compressed into a virtual and magic spectacle. The new concept after elaboration of the blend is a “powerful” and “magical” genie application serving at your wish in a “spectacular” multitasking environment. The following diagram illustrates this immediate blend.

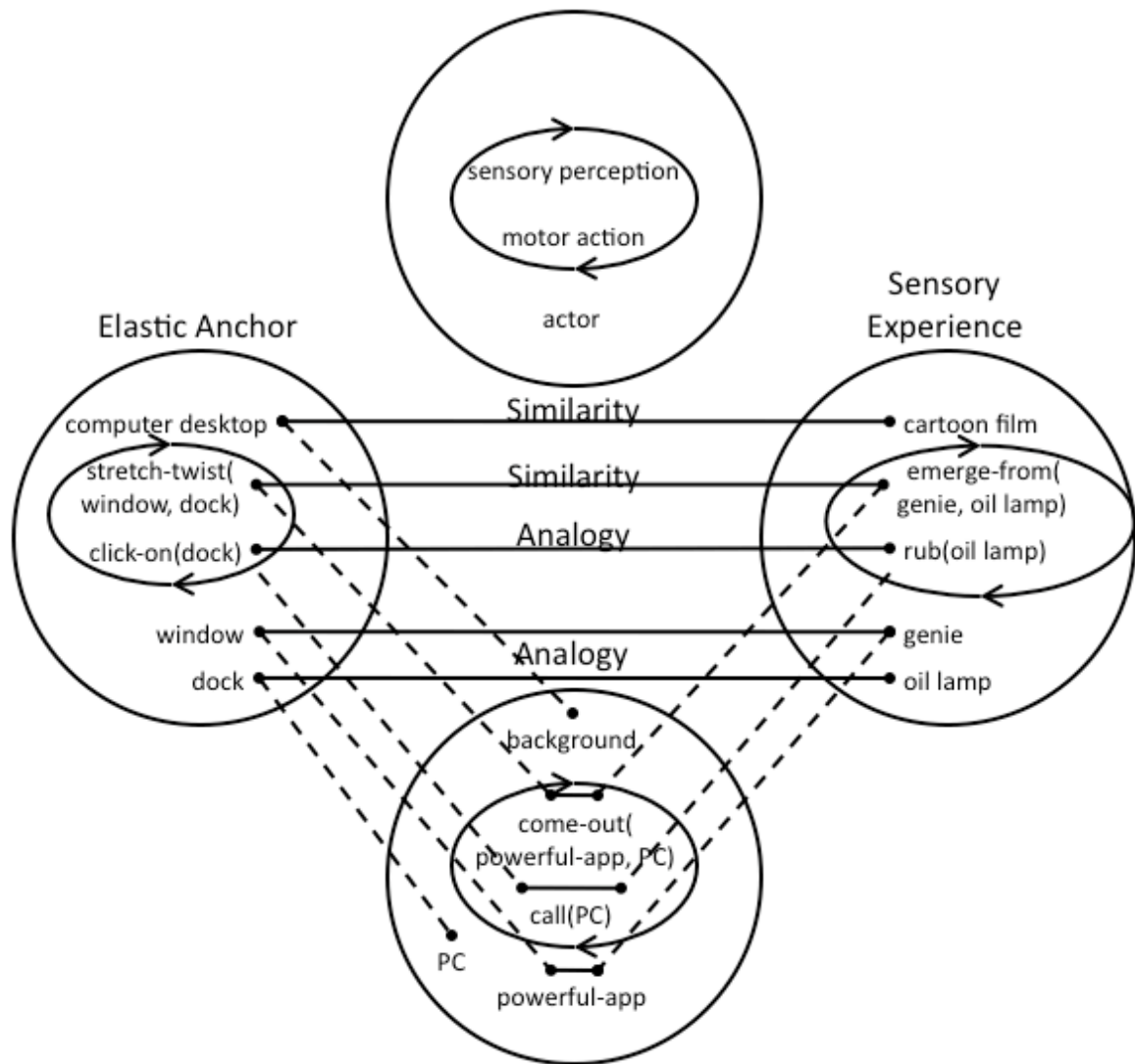


Figure 6.12 The immediate blend in the genie effect of Mac OS X

Bodily Engagement

Since Sutherland's *Sketchpad* and Engelbart's computer mouse demonstrated the spatial manipulation feature of pointing devices, the mouse, the stylus and tablet, or lately the touchpad, has broadened the horizon of motor input other than the classical command-response mechanism. This paradigm draws our attention to the dimension of space and to the least extent, some parameters of motion, like direction and distance. For instance, the layout of the desktop always depends on which direction and how far we drag the icons. The dock is also usually set to show when the pointer moves to a

particular far edge of the desktop. However, other specific qualities like moving speed or motion path do not really matter much in most cases. Bouncing of a clicked icon does not depend on how the pointer reaches it. That is to say, the motor input to the GUI is not very motion-sensitive. In addition, the Macintosh OS X environment seldom shows continuous and autonomous transformation without user input, except the start of screensaver programs. Although some screensavers do show visual effects playfully altering the appearance of the desktop, these effects do not actually change the desktop arrangement because everything would be reverted once the user activate the machine again. Hence, the engagement in the GUI is largely halted.

Liveness

It is quite an alternative view to see the use of an operating system through the GUI as a performance. As I have argued in Chapter 4, today's operating systems, like other real-time control systems, turn users into puppeteers, as well as audiences, of the performance animation enacted by those widgets. In the Macintosh OS X environment, users can trigger and then watch the genie-like performance of windows, or orchestrate and gaze at the harmonic wave sliding across the dock like a line of dancers. In these visual effects, the stored visual materials should include all the graphical icons and the window motif. Meanwhile, the scaling of icons and the stretch and twist of windows are rendered in real time. The degree of liveness is fairly live, indicating that the GUI environment is quite contingent and open to user intervention at all times.

The Water-Level Interface

(Type of liveliness: Emergent; Level of understanding: Immediate and Metaphorical;
Mechanism of engagement: Continuing; Degree of liveness: Highly live)

While the term user interface primarily refer to computer interfaces, today many electronic devices and gadgets also have user interfaces. Compared with the

homogeneities like the desktop metaphor in the GUI of personal computers, the designs of gadget interfaces are much diverse because the uses of these devices contrastingly differ from one another. For instance, a touch-screen interface might be convenient for a digital camera user to select a region to be focused, but might not work very conveniently for a driving cell phone user who wants to rely on the tactile sense to make a dial while keeping the visual sense on the road. However, the variety in the context of use also opens up possibilities of novel designs. An example is the aforementioned “water-level” interface of the NEC mobile phone FOMA N702iS. The intriguing gestural interface displays computer-generated imagery of liquid water that looks as if “contained” inside the phone’s screen. When the user tilts the phone, the rendered water surface moves in response to the action. Moreover, the water level represents the battery level. In other words, the interface is also a running battery meter.

Liveliness

The perceptually realistic water wave effect displayed in the interface constitutes an emergent type of liveliness. Although the water wave movement is reactive to the user’s hand action, the resulting interactional liveliness is very least because the action and reaction is unidirectional only. That means the water imagery is always passive. Conversely, the wave presents a rhythmic kind of patterns that is perceptually reversible, especially when the user keeps shaking the phone. The resulting successive wave effect is emergently lively.

Understanding

The simple water simulation in the interface surprisingly invokes multiple levels of understanding in the user’s mind. As described in Chapter 4, the first blend takes place immediately and effortlessly. The subtle water movement first incites the user to slightly tilt the mobile phone, which in turn senses the user-driven motion by means of the built-

in accelerometer and computes and displays the reaction of the water surface. This motor-sensory loop constitutes an elastic anchor, blending with the user's sensorimotor experience of holding a bottle of water, yielding an imaginative conceptualization of a water-filled mobile phone. The representation relation between the water imagery and actual liquid water is tightly compressed into a virtual form of water (which differs than real water at least in mass), and the analogy relation between the phone and the bottle is compressed into a new concept of container-phone. The following diagram illustrates this imaginative conceptualization (reprinted from Chapter 4).

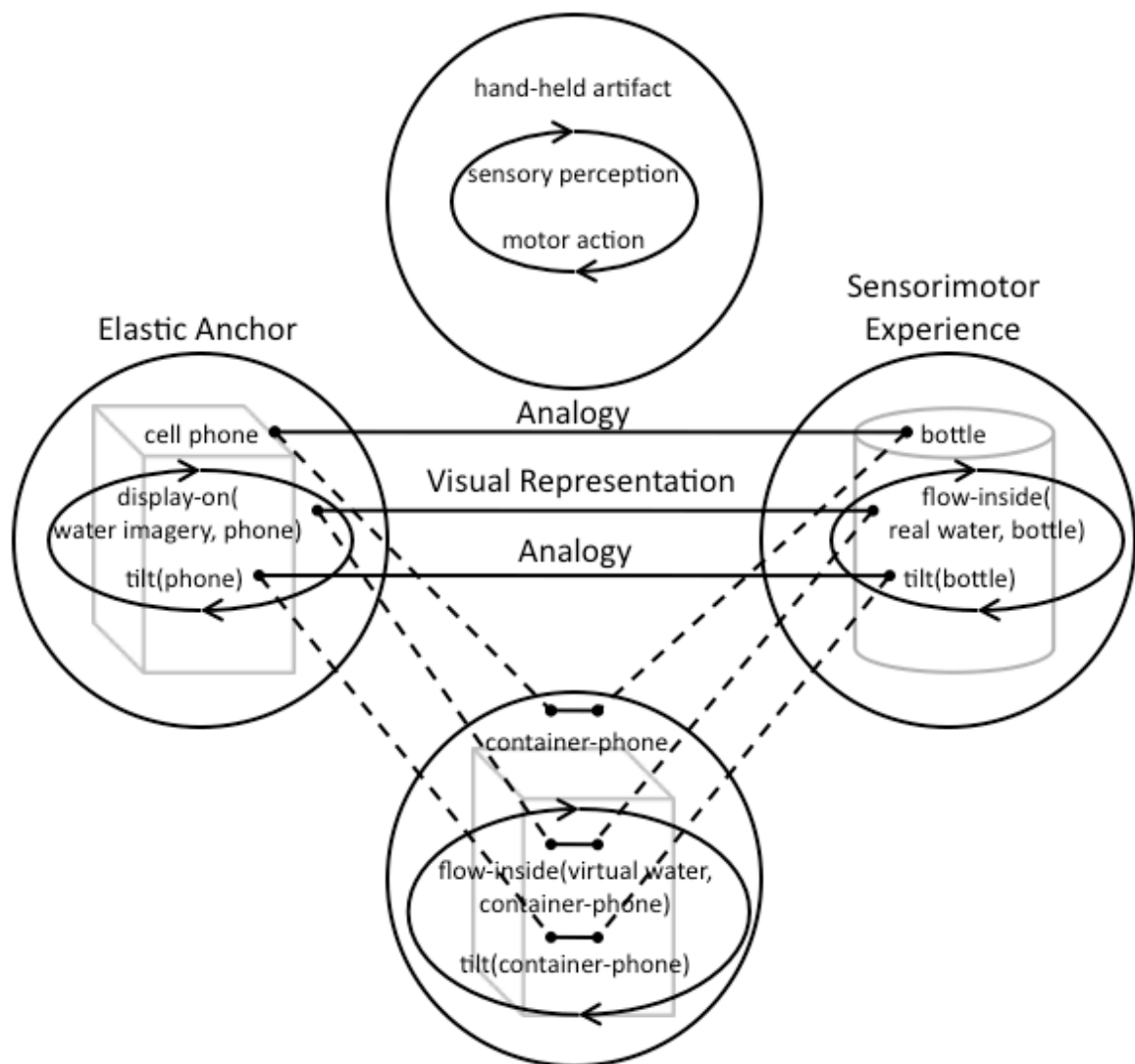


Figure 6.13 The immediate blend in the water-level interface

The secondary blend emerges at the metaphorical level, when the user notices that the water level is descending gradually and reflecting the ongoing energy consumption in the container-phone. This articulation gives a concept of a level-type indicator, which is shared by the interface and an ordinary battery meter. The analogy relation between the act of checking water level and checking battery meter level is compressed into checking how much resource is left. The frame of conservation is invoked. The blend elaborates an imaginative narrative about someone in some severe environment monitoring some limited resource, leaving a provocative message: “save the juice!”

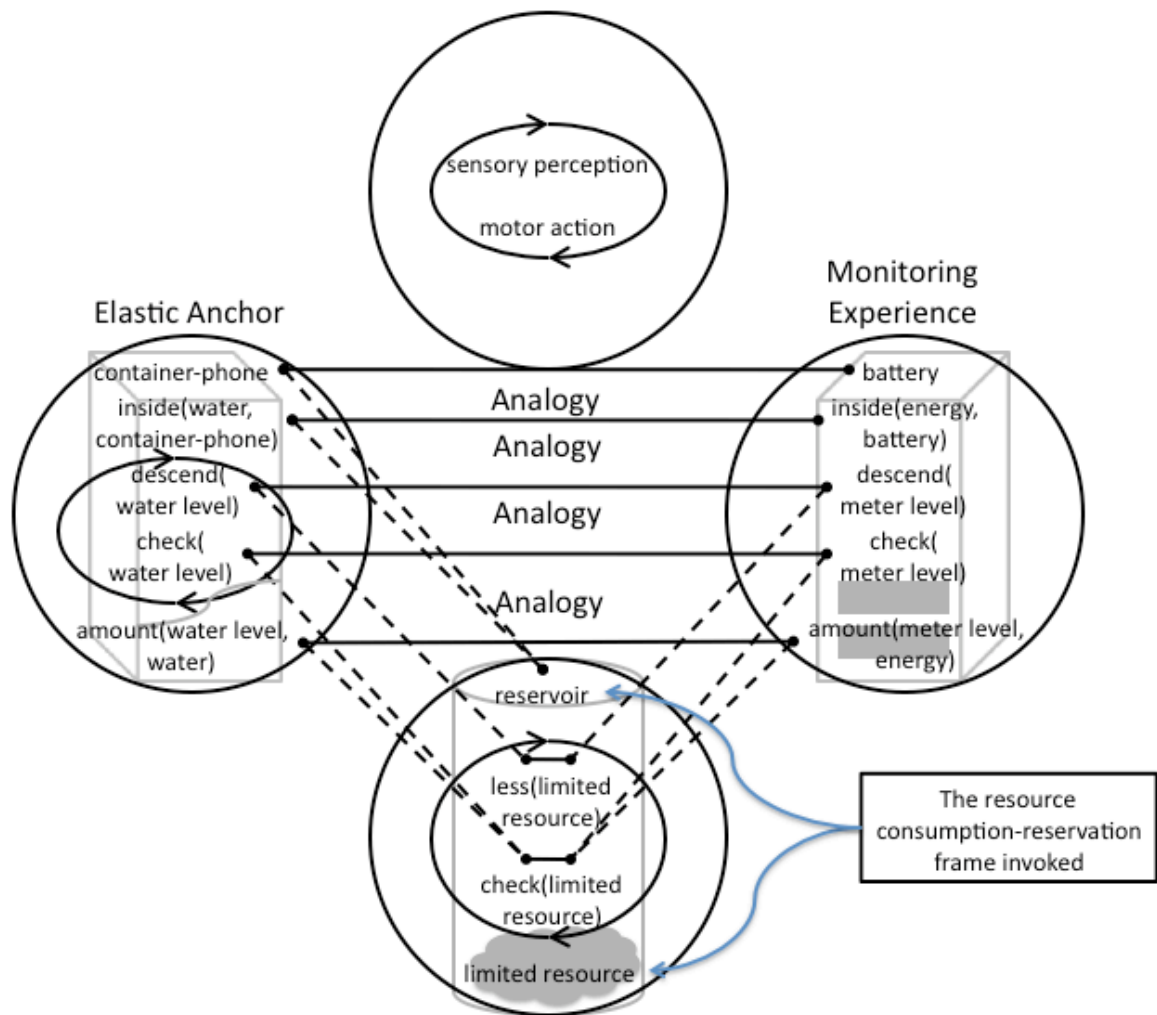


Figure 6.14 The metaphorical blend in the water-level interface

The water movement shown in the interface invites the user to shake the phone. The resulting motion will be accurately measured by the built-in accelerometer, and the responding water wave will be simulated and displayed accordingly. That is, the changes in speed and orientation of the user's motor action determine how the virtual water surface reacts. The motor input to the interface is obviously motion-sensitive. Moreover, the water level is persistently descending with or without user input, as well as divergently waving with respect to user shake at any particular level, in which water waves at higher levels of water definitely flow very differently from those at lower levels with less water. Hence, the interface environment is enduring. To sum up, the interface entails a continuing engagement.

Liveness

Like the GUI of personal computers, mobile device interfaces are generally not related to any sense of performance. First, they are designed for individuals, a very personal artifact. Second, they aim at high productivity, rather than amusement. However, after seeing a series of engaging and focusing seminars by Steve Jobs with his popular "companion," the iPhone, on stage, it is completely sensible to say that interaction with a user interface can be a form of performance. The FOMA N702iS invites its users to take action in the performance, and engages them with moving images, which are totally rendered instantaneously. There is nearly no stored visual content in the interface. Hence, the interface is very live, echoing the contingency of life.

The interfaces discussed in this section can be regarded as utilities applications of which the primary objective is about productivity and usability. They are designed for certain functional purposes, such as measuring, organizing, and presenting information. Yet, they perform quite differently in terms of the qualitative variables. It follows that generative and interactive animation is orthogonal to the practicality of an artifact.

Multimedia Websites

Apart from amusement and utilities, generative and interactive animation also plays a substantial role in mass communication, mainly over the Internet. The Internet was originally a communication network for specific communities, like academia. Yet, in late nineties of the last century, the World Wide Web has made the Internet emerge as the massive channel of general communication. From the text-based HTML-encoded websites of the early days, to today's ubiquitous flash-enabled multimedia websites, animation has become an integral part of presentation, and also representation, on the Internet. Because of animation, many multimedia websites are not only able to present information but also successful in representing illusory virtual environments that absorb users into imaginative worlds and stories. This section talks about two examples, one of which is a popular and addictive website intended for personal messaging, and the other is a typical one for promotion.

SnowDays by Popular Front

(Type of liveliness: Emergent; Level of understanding: Immediate and Metaphorical;
Mechanism of engagement: Continuing; Degree of liveness: Live)

Messaging, besides web browsing, is another major activity on the Internet, and quite many websites have incorporated the feature of messaging. For example, one could easily find on Amazon.com a hyperlink labeled "Share with Friends," which allows one to send friends electronic recommendation of an item. While many of these sites make messaging an explicit option, *SnowDays* (mentioned previously) incorporates this feature more seamlessly and elegantly in its virtual world. The popular website is a greeting web page presented by the digital service agency Popular Front. It depicts a scene of a snowing day in which every falling flake was "handcrafted" by a web visitor. The visitor may leave a message "inside" the flake and the receiver would be notified by the website. The receiver, or in fact any other visitor, may go to the page, look at snowing, appreciate the meticulous details of the falling flake, and respond. In other words, the site projects

visitors an illusion of messaging through snowflakes in the virtual world (Figure 6.15). Since the launch in 2002, the website has accumulated over nine million user-created snowflakes by the end of 2009.



Figure 6.15 Cut-out snowflakes tailor-made by individual users carry their greeting messages evocative of personal memories

Liveliness

The page of *SnowDays* displays a graphical representation of an outdoor open space, with a few trees, covered with snow. Snowflakes of different sizes, some of which are large enough for the feathers to be seen, are descending constantly and steadily all over the screen. The movements of these flakes do not seem to be driven by a particular impulse, nor do they result in any irreversible consequence in the environment (for example, the snow on the ground will never get thicker). In fact, the falling flakes do not

really land on the ground, but loop back to fall from the sky again sooner or later. Although this recurrence of snow differs from actual snowing, the rhythmic effect in *SnowDays* still imbues the virtual snowy space with emergent liveliness.

Understanding

The recurrent snowing images of the website is able to constitute multiple levels of understanding. When a visitor glazes at the graphical scene and moves the pointer over any particular falling graphic, a close-up view of a snowflake pops up and a motor-sensory immediate blend comes about. The representation relation between the falling sprite and an actual falling flake are compressed into a piece of virtual flake in delicate sixfold symmetry. The motor action of moving the mouse, or a finger in case of the touchpad, in the elastic anchor connects to the act of catching a snowflake to examine in a real snowing day from one's past sensorimotor experience. Similarly, the perception of zooming connects to that of looking closely at something. These links are compressed into an imaginative act of pointing for examining a virtual flake when one is looking through a "window" in a snowing day. This immediate imaginative blend is illustrated in the following diagram.

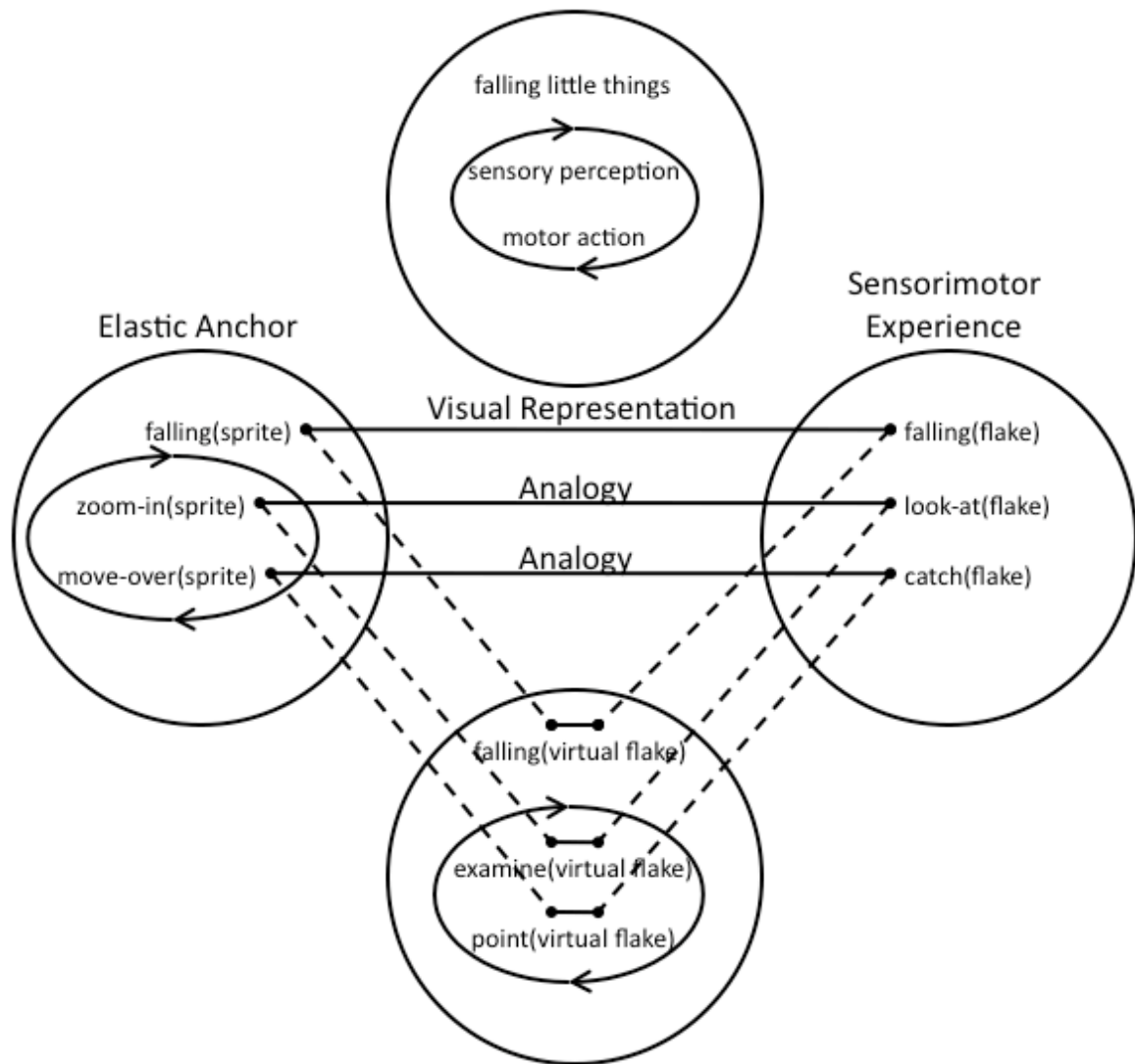


Figure 6.16 The immediate blend in *SnowDays*

Apart from exhibiting the beauty of ice crystals, the virtual flakes in *SnowDays* have other even more intriguing features. First, each flake carries a message from its creator to the receiver. This messaging function makes the flake comparable to airmail, a long-distance call, or even a note from the sky. The website narrows the distance between the sender and the receiver. Second, the sender creates a snowflake tailor-made for the receiver, with the tiny tool provided in *SnowDays*. The tool simulates the little trick that many people learned to construct their cutout snowflakes in elementary schools. The memory of such experiences can be integrated with the sensorimotor experience of using the site, and then compressed into an act of making a gift for the special one, with the

season's greetings frame invoked. This metaphorical blend elaborates an imaginative narrative of exchanging heartfelt greetings and handcrafted gifts with someone from a distance in a lonely winter day. With the attendant visceral and nostalgic association triggered, the hidden meaning is evocative and affective: “you are not alone!”

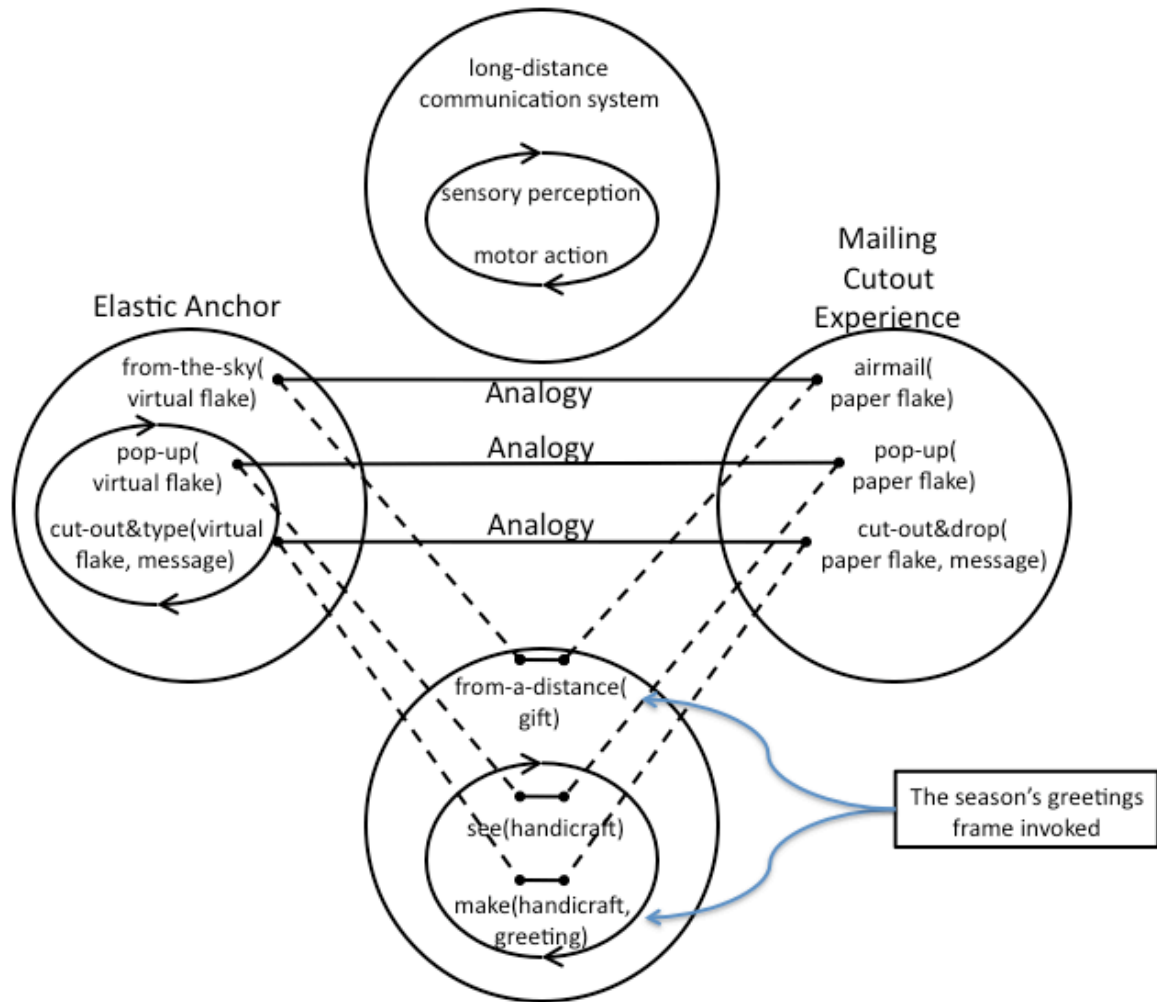


Figure 6.17 The metaphorical blend in *SnowDays*

Bodily Engagement

The sensorimotor experience of using the website relies not only on the animated graphics it presents, but also the bodily engagement it affords. When first being shown the snowy scene on the website, a visitor is invited to catch a falling flake. One can move the pointer freely to pick a flake, but timing and motion path are very determining. Since

all flakes are descending continuously, the visitor has to perform well with eye-hand coordination in moving the pointer to reach a flake in motion. If the visitor misses the time, the flake would fall out of frame and might not come back in the near future. If the visitor moves the pointer toward a target too hastily, he might run across an unintended flake in the middle of the path and zoom in another flake. Moreover, the cutout tool of the site requires a visitor even more attention to the direction and distance of the pointer motion, because it is just like how we do cutout by using scissors. All these embodied interaction mechanisms show that the site takes motion-based motor input.

In addition, *SnowDays* also demonstrates an enduring environment. Although the virtual snow on the ground, as mentioned, would not get any thicker as time goes by, the scene does show other enduring transformation to reflect the change of time. With the Internet connection, *SnowDays* is able to change the background color and atmosphere according to the actual time of day and weather. For instance, the site would show an orange sky during sunset. Furthermore, with the unlimited number of web surfers (most of them just want to kill time) at all times, there are always new flakes falling into different parts of the scene even though a user just gazes into the scene without taking any action. Therefore, the virtual snowy space is enduringly changing, and the mechanism of engagement is continuing.

Liveness

The *SnowDays* website enables its visitors to create their own flakes, packed with messages, and then let the flakes descend in the virtual world. The visitors act like wizards in this imaginative world, collectively manipulating the virtual snowing. Meanwhile, they are also individual audiences of the beautiful natural scenery. In short, what *SnowDays* presents is an improvised performance between the site and the visitors. In this performance, the major visual components include all the graphic sprites. With over millions of flakes, the amount of sprites should also be huge. Fortunately, all flakes

are sixfold symmetric. The system only needs to store the cutout path, and a complete sprite can be generated by simple procedures. In other words, less than one-sixth of graphic content is really stored, together with other static background elements. Since all the snowing and changing sky color effects are also generated in real time, the resulting performance can be regarded as live.

Come Clean

(Type of liveliness: Intentional; Level of understanding: Immediate and Metaphorical; Mechanism of engagement: Halted; Degree of liveness: Fairly live)

Besides messaging and social networking, the Web is also a common platform for promotion, marketing, or advertising today. Rather than the conventional push strategy used in old media like television or newspaper, many web-marketing campaigns adopt the pull strategy, motivating consumers to request information, services, or products through the Internet. With the bidirectional communication channel and interactive multimedia websites, advertisers are able to immerse potential consumers in imaginative worlds or narratives, arousing their interests in some intended promotional items.

Comeclean.com is the website of a typical promotional campaign. The site invokes standard interface mechanisms like data entry and real-time composite moving images to make web visitors believe that the advertised product can help wash away sins. A visitor is first presented with a subjective view of looking down at a sink basin and is welcomed by a consoling voice speaking in English with a South Asian accent, prompting a confessed message. Once the visitor types in a message, the site shows that the message is handwritten on a lady's palm and then is washed away with cleansing foam. Moreover, the visitor can see confessed messages from other online users also being cleaned up. Both personal and social experiences persuade one to look into the advertised cleaning supplies (please see Figure 4.19).

Liveliness

When a user first enters the website, one hears the consoling voice and the sound of tap water. Yet, the photographic image is static, until the user types in a message. To the user's surprise, the lady's palm has the message handwritten on it! This prominent act of showing is an intentional type of liveliness. The user wonders what she is going to do next. It follows that she applies the cleansing foam and washes the dirt away. Throughout the whole process, the lady's prominent action leads the user's attention and provokes him or her into taking the intentional stance.

Understanding

Like most other advertising campaigns, Comeclean.com provokes multiple levels of understanding. At the immediate level, it entails a motor-sensory connection between keyboard entry and real-time composite imagery. Typing in a message and then seeing it on the lady's palm constitute an elastic anchor, blending with a user's sensorimotor experience of writing notes on the palm. The typing action in the elastic anchor has an analogy counterpart, the action of writing notes, in another input space. The identity relation between the lady's palm and the user's palm is compressed into a virtual palm of the user, which shows whatever the user intends to "write" on it through the keyboard. Meanwhile, the real-time composite moving images represent what the user would see on the virtual palm. This blend can be shown in the following diagram.

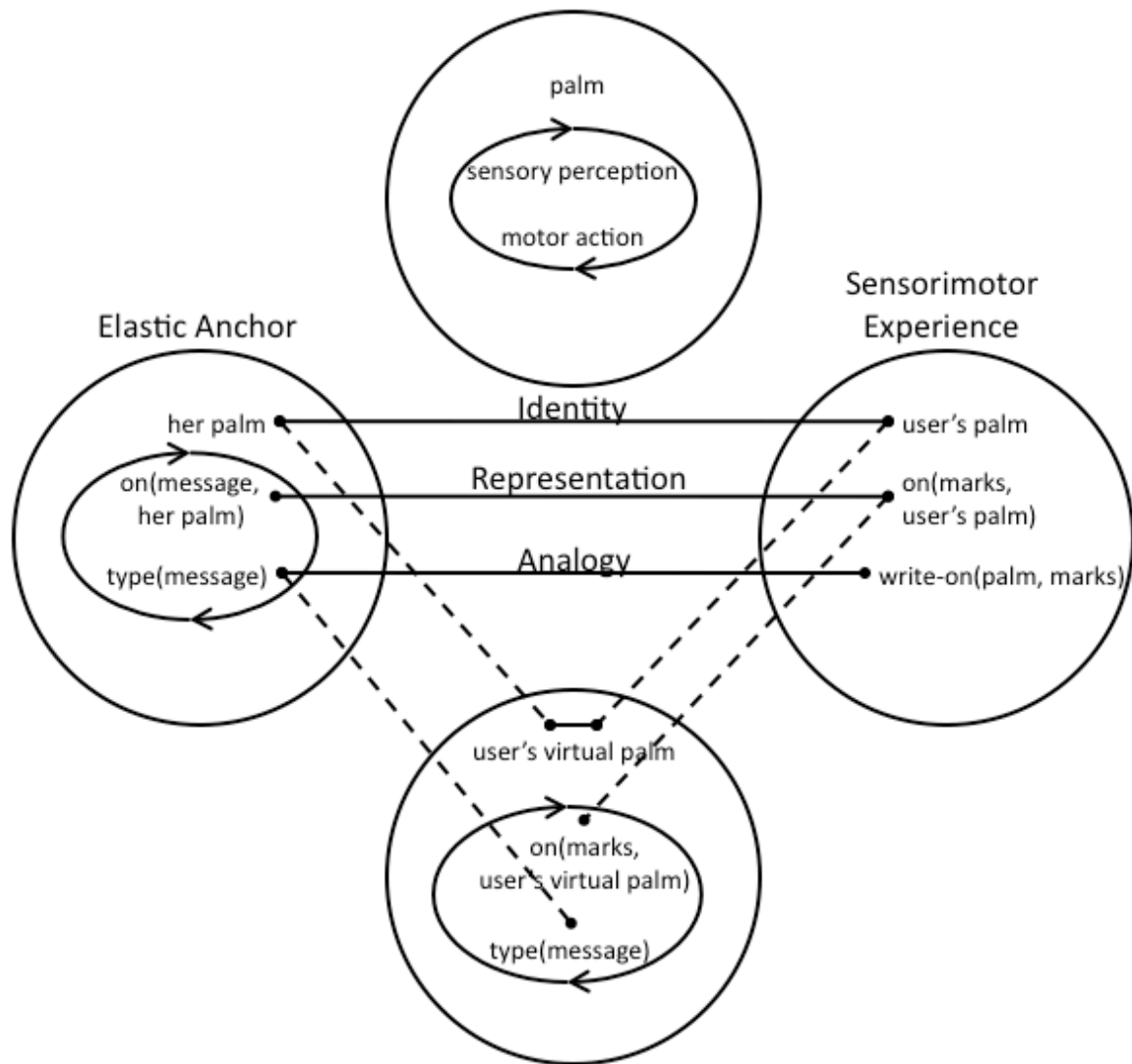


Figure 6.18 The immediate blend in Comeclean.com

At the metaphorical level, the greetings made by the consoling voice in South Asian accent invoke the confession frame, making the user form a metaphorical blend between the micro-narrative of washing one's hands using particular cleansing supplies and the user's wish image of becoming spiritually clean after confessions. The analogy relation between the consoling voice and the priest during confession is compressed into a conceptualization of spiritual-physical cleansing foam. The output of the blend is an imaginative image of the foam that is able to "clean up" both the body and the soul, motivating the user to look into the product.

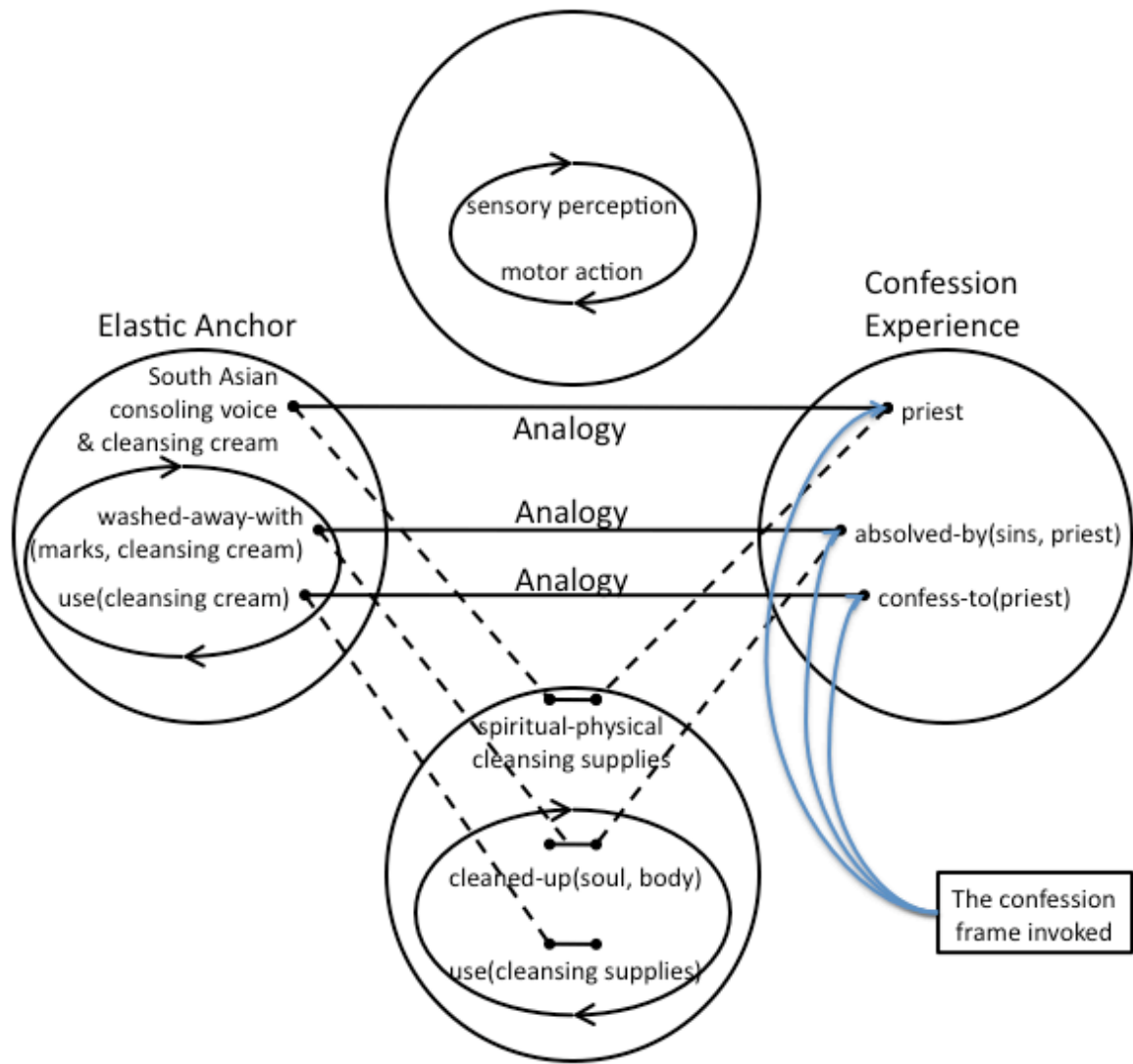


Figure 6.19 The metaphorical blend in Comeclean.com

Bodily Engagement

In Comeclean.com, the major motor input is done through a classical input device, the keyboard. A user is prompted to type a message. Before the user hits the return key, the website keeps displaying the subjective view of the basin with water constantly flows from the tap, meanwhile ambient sound and music come along ceaselessly. In fact, both the visual and audio are looping. The website would wait indefinitely for user input. Once the site receives an input, it responds with a real-time composite movie, and then waits for another input again. This form of alternate input-

feedback sections is a typical halted bodily engagement in interactive digital environments. Moreover, unlike *SnowDays*, which continuously shows falling flakes created by other online users while waiting for user action, Comeclean.com requires the user to explicitly switch to another page for viewing other people's confessions. This switch further disrupts the continuity, that is to say, "becoming," of the environment.

To make the engagement more continuing, suggestions include replacing keyboard input by "finger-writing" on the touchpad enabled by character recognition technology, and showing others' confessed messages intermittently while waiting for user input. The former measure enables more motion-sensitive input, and the latter projects an illusion of an enduring environment.

Liveness

The website provides a platform on which the user and the system co-perform a ritual of confession. The user is responsible for the message input, while the system takes care of generating photorealistic visualization. This consoling ritual is only fairly live, because most of the presentation materials are pre-recorded and stored, including photographic moving images, voice-over, and the music. The only part that is open for user control is what is written on the palm.

The Web is an emergently massive platform for different forms of social communication, like personal messaging and online marketing. This section introduced two examples corresponding to the suggested forms. They are selected because they demonstrate the principles of the new animation paradigm. Compared with the ordinary use of traditional film animation as primarily mass propaganda, generative and interactive animation is able to permeate particular social networks, triggering livelier communication through the Internet.

Digital Signs

Regarding the role in social communication, animation first emerged in the old media like the cinema, and then pervaded the Web, and now it spreads to even the physical communal space. When technologies of digital displays or projections are becoming more accessible today, animation phenomena can be found in many of our digitally mediated public areas, such as on an LED sign in a train station, on an information kiosk in a shopping mall, on the façade of a building, and even on the traffic lights of a pedestrian crossing. This section talks about a very inspirational and innovative use of animation on traffic lights to invigorate communication.

Animated Traffic Lights

(Type of liveliness: Intentional and Emergent; Level of understanding: Immediate; Mechanism of engagement: not applicable; Degree of liveness: Slightly live)

Traffic lights for pedestrians conventionally use static icons of “green man” and “red man” alternately to signal people when to cross the street. Some exceptional traffic lights in a few cities, for example, in Taiwan, Spain, and Mexico, display LED-lit animated green men to tell pedestrians *how* to cross the road. During each green signal, the little green man walks in a cycle. After a while, he starts running faster and faster as the system counts down to the red signal. This animation phenomenon, like those in animation films, is considered to be non-interactive, because the system does not take a pedestrian’s bodily response into account when generating subsequent output. Hence, I exclude the discussion on the bodily engagement of this artifact here. Besides, all other three variables, including the amount of stored component in playback, are worth closer reading.

Liveliness

The little green men in these traffic lights demonstrate emergent liveliness as well as intentional liveliness consecutively. First, when a green man is walking rhythmically

and steadily in a cycle, his repeating action does not seem to bring about any irreversible consequence. This walk cycle is an emergent pattern. Second, when the green man suddenly starts to run, it seems to a pedestrian that the green man has initiated this prominent action. One cannot help taking the intentional stance and wondering what the green man is running for. It is probable that dangers are pending. Here, the intentional liveliness draws a walker's attention from a relatively diluted mode (just walking steadily) to a more concentrated mode (better hurry up).

Understanding

When seeing the little green starts to run, a walker would immediately feel that the green man is in hurry because the perceived motion evokes one's motor experience in case of a rush, as the research of mirror neurons has suggested. The walker would then blend the green man's situation with his or her own situation of crossing the road. In the blend, the outer-space link between the green man and pedestrians is compressed into an inner-space identity relation, which implies that the green man stands for all walkers, including the viewer. The traffic light is linked metonymically to the road. Seeing the green man running on the traffic light integrates with the viewing walker's experience of seeing other human walkers running across the road. Elaboration of this immediate blend outputs an image of running with other imaginative walkers on the road, compellingly suggesting the viewing walker: "hurry up!"

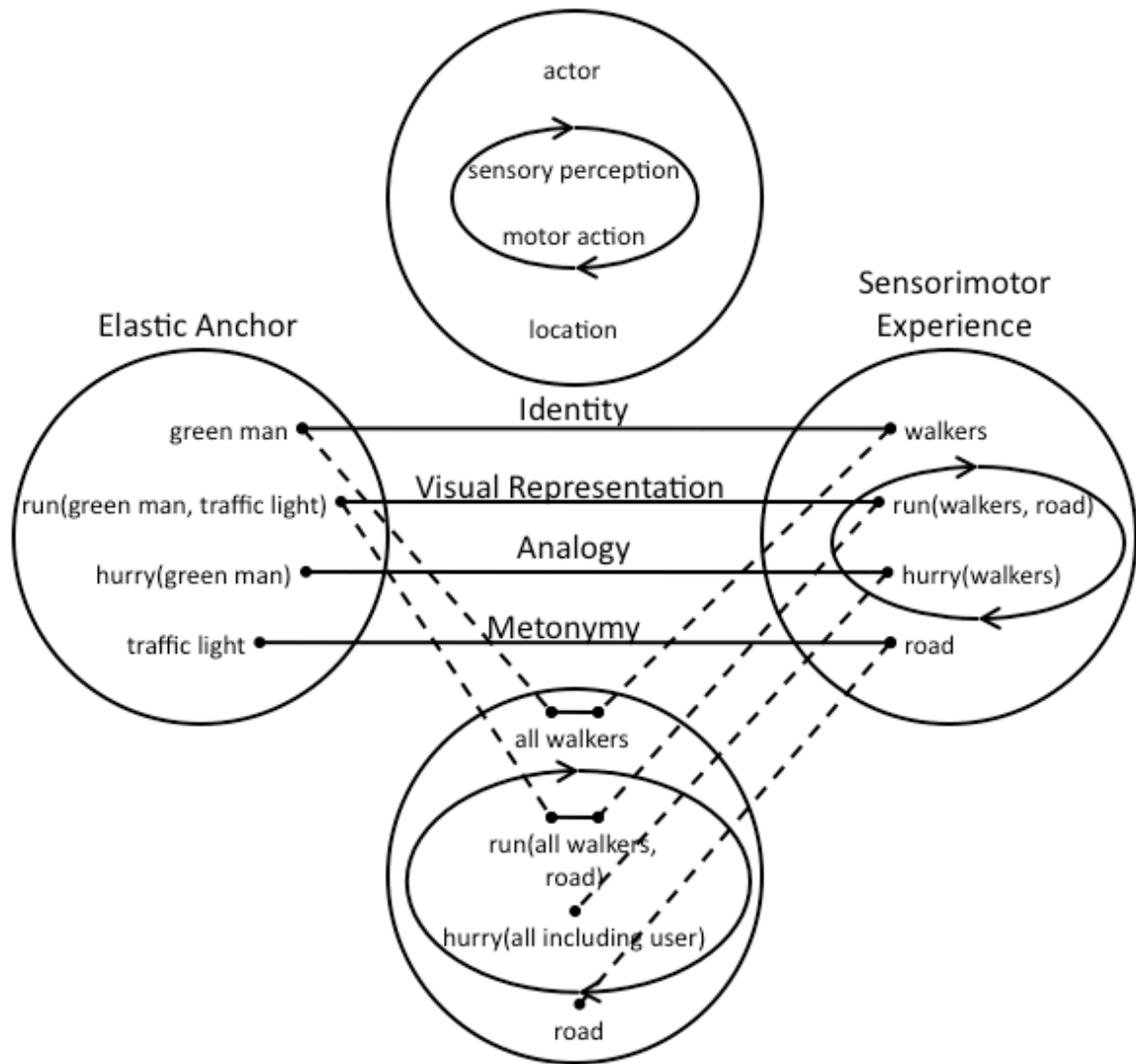


Figure 6.20 The immediate blend in traffic lights

Liveness

The animated traffic lights incite pedestrians' bodily responses. On seeing an animated green man on a traffic light, a walker would either keep walking or accelerate. Although the traffic light system does not take into consideration these responses to output next walking cycle, the artifact differs from simple playback of stored animation videos in that it is able to drive viewers to take action. To a very least extent, all pedestrians are supposed to "perform" (walk or run) with the little green man. In this co-performance, the durations of walking and running vary in different crossings. This

divergence in timing is made possible because the animation of the green man is not a purely playback of stored content. In fact, the stored visual material is only a looped sequence of rudimentary images in a matrix of LED points representing a walk cycle. The illusion of changing from walking to running is created through altering the playback speed. Therefore, the pace of green men on different traffic lights can be programmed and adapted to different traffic conditions accordingly. This subtle variation might not be noticeable to most pedestrians, but it does make the system slightly live.

With more and more electronic displays or projections being installed in our everyday communal areas, we can actually see the new animation phenomena not only in living rooms, classrooms, or offices, but also in café, buses, or even on the street. Although the traffic light system described in this section is not interactive, it opens up new possibilities of creating an expanded illusion of life in communal areas. When some of these displays become touch-sensitive, or when projections come with motion-sensing technologies, I believe animated artifacts in the communal space will be like their counterparts in digital environments, becoming reactive to people's bodily action and sensitive to their motion in the near future.

Digital Artworks

After discussing such domains as amusement, utilities, and communication, this section looks at artifacts in the realm of the arts. Art has been a platform for humans to express, to interrogate, and to experiment. It contrasts with all the previous domains in terms of creators' motivations, but it also shares with them in demonstrating certain extent of animation phenomena. Many works of art, like Alexander Calder's mobile sculptures, mechanical automata, not to mention animated films, are exemplars showing an illusion of life. For an expanded illusion of life, digital artworks, by which I mean works whose discourse processes rely on the use of digital technology, are particularly illustrative. This kind of works usually incorporates computer programs to produce

variable instances that show autonomous, reactive, transformative, and contingent behaviors. Among the innumerable digital artworks that have emerged in recent decades, I select one canonical and celebrated installation work, Camille Utterback and Romy Achituv's *Text Rain* (1999), and one more peculiar but also well-known interactive web comics, Han Hoogerbrugge's *Modern Living* (1998-2001). Although the former is an interactive video installation and the latter purely runs on the Web, I see both of them as digital artworks because I found that generativity and interactivity are indispensable in their meaning-making processes.

Camille Utterback and Romy Achituv's *Text Rain*

(Type of liveliness: Intentional, Interactional, and Emergent; Level of understanding: Immediate and Metaphorical; Mechanism of engagement: Continuing; Degree of liveness: Highly live)

Text Rain is probably one of the most widely known works of digital art by the end of the last century. The award-winning work has been exhibited in many art festivals around the globe, and was collected in several art museums. It has also been widely covered and closely discussed in the literature of digital art, including Jay David Bolter and Diane Gromala's *Windows and Mirrors: Interaction Design, Digital Art, and the Myth of Transparency* (2003). I do not intend to re-celebrate its vitality with respect to art theories and practices. Instead, I focus on how this artwork manifests the new animation phenomena in terms of the variables.

Liveliness

The work shows a projection of animated letters falling like rain. A participant standing in front of the screen is able to see her or his mirror image being projected in the virtual rain (Please see Figure 4.17). The mirror image is in fact created by flipping horizontally a real-time image captured through a camera placed by the side of the

screen. When the participant raises hand, she would see her screen image following, exactly like what she sees in a mirror. Hence, one can “move” her screen image to catch, lift, and then let fall any letters in the rain. The overall moving images on the screen project a holistic illusion of life. First, the letters fall continually and randomly all over the screen. There is no single obvious impulse driving the falling letters, and the virtual rain does not seem to result in any progression of events or happenings. It is an emergent type of liveliness. Second, when a participant moves her mirror image in the rain, letters might bounce off and slide away from her body. An observer would take the physical stance to perceive this interactional liveliness. Third, each raindrop is in fact a letter coming from a poem, and sometimes a participant might want to look for a word or a phrase along the body silhouette, resulting in abrupt movement or bizarre posture. This kind of behaviors would provoke an observer into taking the intentional stance, looking into which word or phrase the participant is targeting. Here, the participant also contributes to the intentional liveliness of the work.

Understanding

As a poetic work of art, *Text Rain* is expressive. It entails both immediate and metaphorical levels of understanding. When a participant moves her body parts and then sees her mirror image following in the rain of letters, one immediately blends this elastic anchor experience with her sensorimotor experience in actual rains (Figure 6.21). The outer-space representation link between falling letters and falling raindrops is compressed into a concept of virtual raindrops. The outer-space link between the mirror image and the participant is compressed into an inner-space identity relation, so seeing the mirror counterpart receiving the rain is the same as oneself being in the rain. The elaboration of this immediate blend results in an imaginary space in which the participant can dance in the virtual rain without getting soaked.

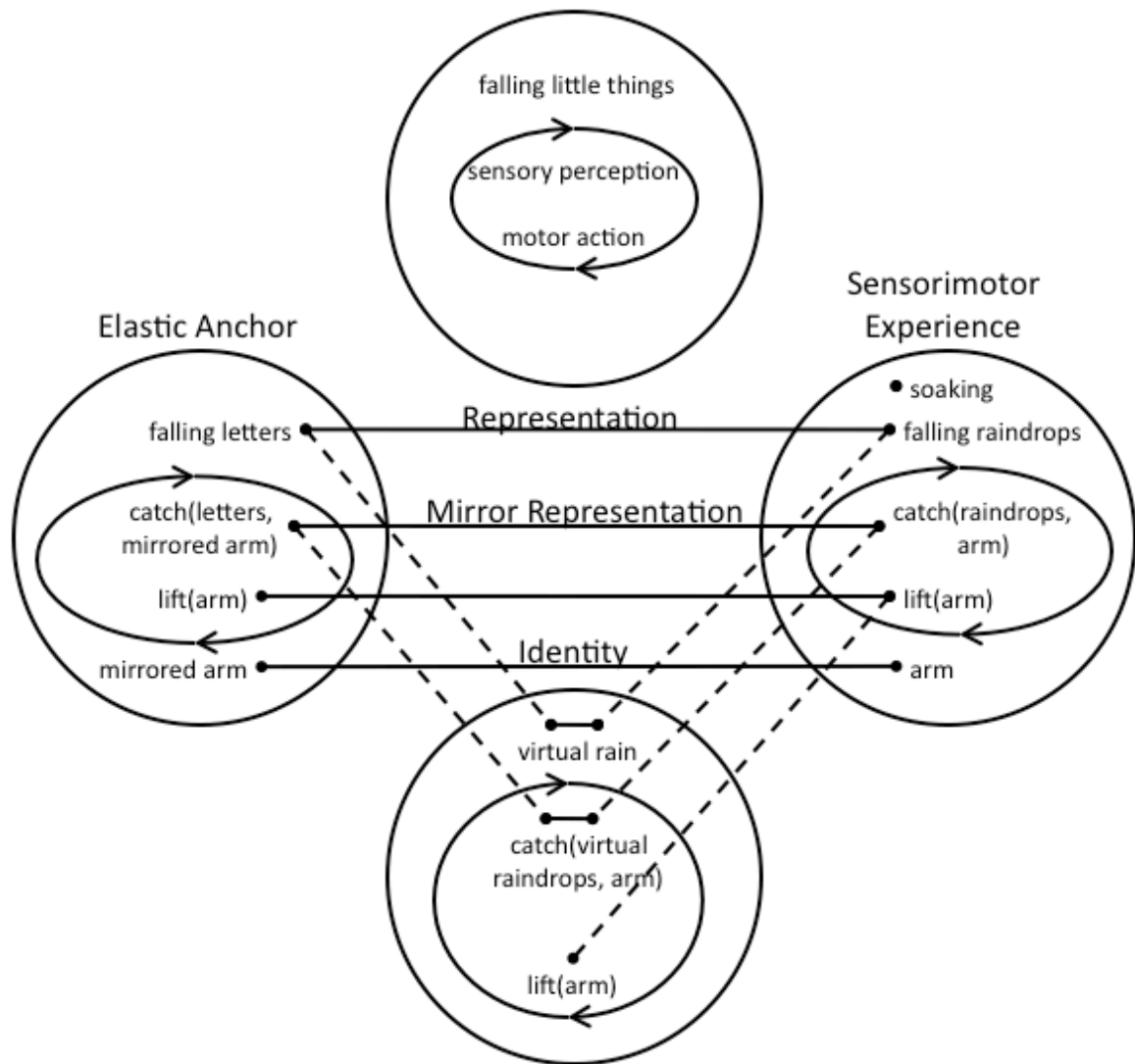


Figure 6.21 The immediate blend in *Text Rain*

The metaphorical understanding emerges as the participant notices that adjacent letters seem to form meaningful words or phrases. Whenever the participant reads the line of letters accumulated along the body silhouette and realizes the resulting verse, the experience is analogical to receiving a celestial message from the sky. The “divine light” frame is invoked and blended with the virtual rain experience (Figure 6.22). The analogy relation between the virtual rain and the divine light is compressed into a new concept of poetic rain, which carries poetic messages. The action of holding raindrops is analogical to the action of embracing light from the sky. Elaboration of the blend provokes a metaphorical narrative of someone like a prophet receiving celestial and spiritual

messages via the body and the surroundings. The animation of falling letters and the interaction with them generate what the artists might describe as an imaginative integration between body, nature, and soul.

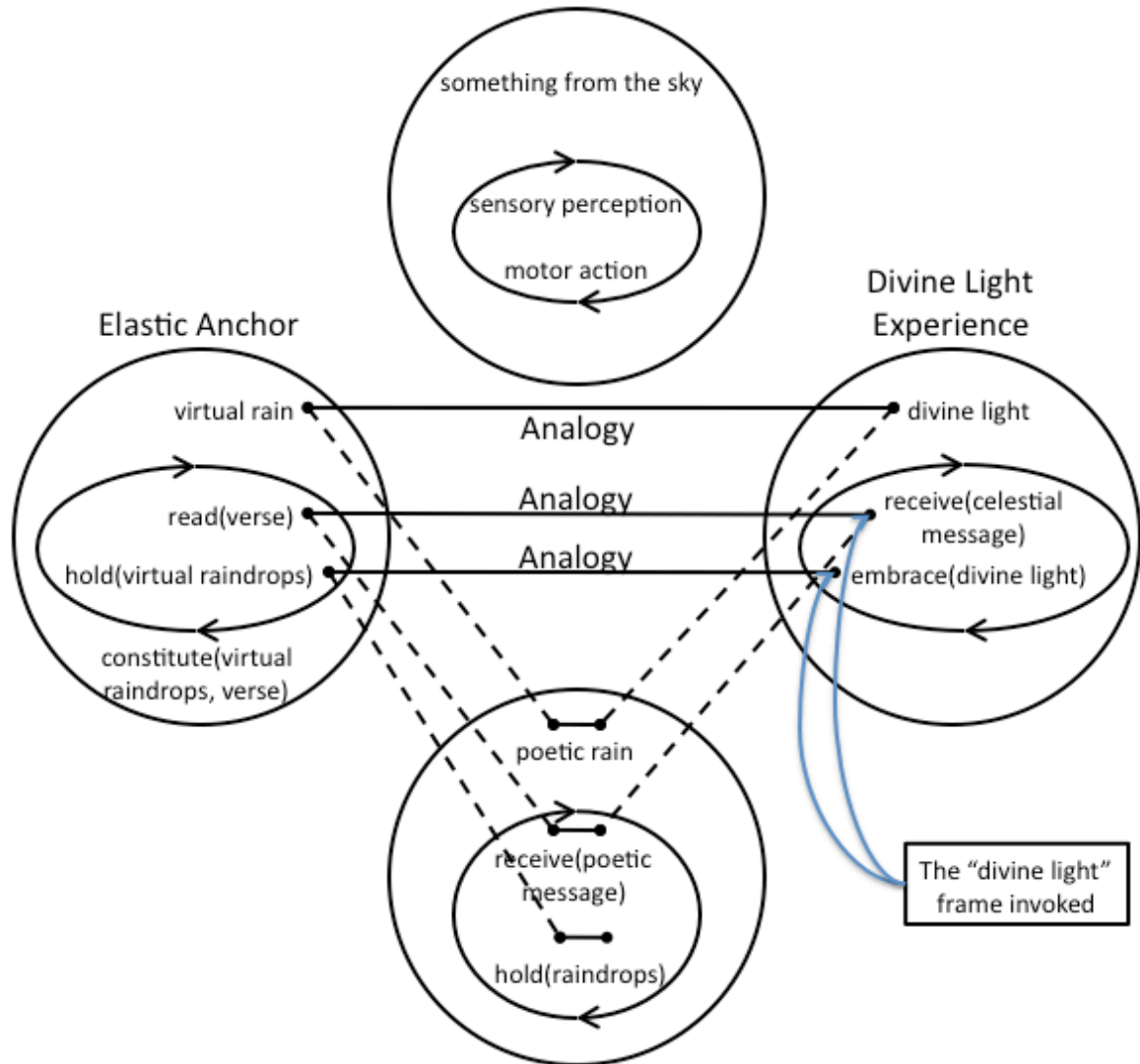


Figure 6.22 The metaphorical blend in *Text Rain*

Bodily Engagement

Making use of real-time video capturing and processing technologies, *Text Rain* is able to accept a user's full-body input without intermediate input device. During this embodied interaction, nearly all motion qualities of motor input are significant. For instance, angle, direction, or speed of a participant's hand movement might determine

which letters are held or slipped away. Both the resulting outlook and message of the accumulated letters would be affected. In other words, the input to *Text Rain* is highly motion-sensitive.

Moreover, the projected environment in *Text Rain* is also enduring. Even without any participant, or the participant doing nothing but just standing still, it keeps falling letters. Since the words from the poem are selected and let fall randomly, the resulting raining pattern on the screen is always transforming. If the participant fails to catch a particular falling verse of the poem, she might not encounter the same verse again near the same position. One has to seek for it somewhere at some other time. This ever-changing environment, together with different acts of individual participant at different time, can lead to very divergent outcomes. All in all, *Text Rain* provides a good example of enduring interaction.

Liveness

Being an interactive video installation that involves the participant's real-time image on the presentation screen, *Text Rain* is intended to be a platform for live improvisation between the system and the participant. As some documentary videos on the work's official website has shown, many participants enjoyed taking part in the performance, with their own creative inputs. For example, someone brought in a real umbrella to perform "dancing in the rain." Several others stood hand in hand and formed the wave to bounce off the rain. They all actively participated in creating the performance animation facilitated by the system. During the show, the only stored visual components are the typographic characters of the poem text. The participant image is captured and reproduced in real time, which would affect the motion of the falling letters on the fly. In other words, the animation is very live, contingent, and open to participant intervention.

Han Hoogerbrugge's Modern Living

(Type of liveliness: Interactional; Level of understanding: Immediate and Metaphorical; Mechanism of engagement: Decelerated; Degree of liveness: Slightly live)

Like many works of art, *Text Rain* was first exhibited publicly in an art gallery. In contrast, another work to be discussed in this section emerged in a strikingly different platform. It was originally intended to show on the Web. Since 1997, the Dutch comic artist Han Hoogerbrugge has started to make use of the Internet publishing interactive animated comics on his website Hoogerbrugge.com. The first work, *Modern Living / Neurotica* (1998-2001), featuring nearly one hundred short animated films, addresses small observations in the artist's everyday life, which also resonate some major issues of our personal life, including workaholics, social impositions, and the like. Apart from a few looped animated GIF sequences, most of the shorts are interactive. The collection can be seen as a documentation of the artist's experiments with mouse-mediated interaction. He is inspirationally successful in remapping conventional mouse actions like pointing and clicking to many different intentions. For instance, in #43 'Itch', a click makes the character (a self-portrait of the artist) "itch" wherever the mouse pointer is located. Through these couplings of motor action and animated feedback, the work mobilizes motor-sensory connection and projects an expanded illusion of life.



Figure 6.23 #43 'Itch' in *Modern Living*

Liveliness

Among the archive of animated pieces in *Modern Living*, we can find at least two types of liveliness, including the interactional and the emergent. The former can be seen

in many pieces in which the character tends to “get away” from the mouse pointer. For example, in #61 ‘Drowning’, a user can move the pointer to play hide-and-seek with the character. In #54 ‘Jumpy’, a mouse-over action seems to drive the character jump around. In #68 ‘Obedience’, keeping the pointer over the character’s head can “bring” him to his knees, and after making a kowtow he immediately stands up again. All these reactions of the character demonstrate interactional liveliness. On the other hand, some other pieces show relatively emergent liveliness by displaying multiple reactive characters that distribute a viewer’s attention over multiple areas of the frame. In #60 ‘New Religion’, the characters line up to form four rows. When the pointer moves across the characters, they stand up and then bend down continuously forming a wave very much like the magnification effect in the dock of Macintosh OS X. In #85 ‘Material Guy’, when a user runs the pointer across the faces of the characters, they swell up and vomit home appliances like washing machines, dryers, toasters, and the like. There are often simultaneous happenings and movements in the frame, which dilute our attention. Moreover, the action and reaction in these pieces are reversible. That is to say, one can reverse the pointer movement to produce similar patterns. Hence, they correspond to emergent liveliness.

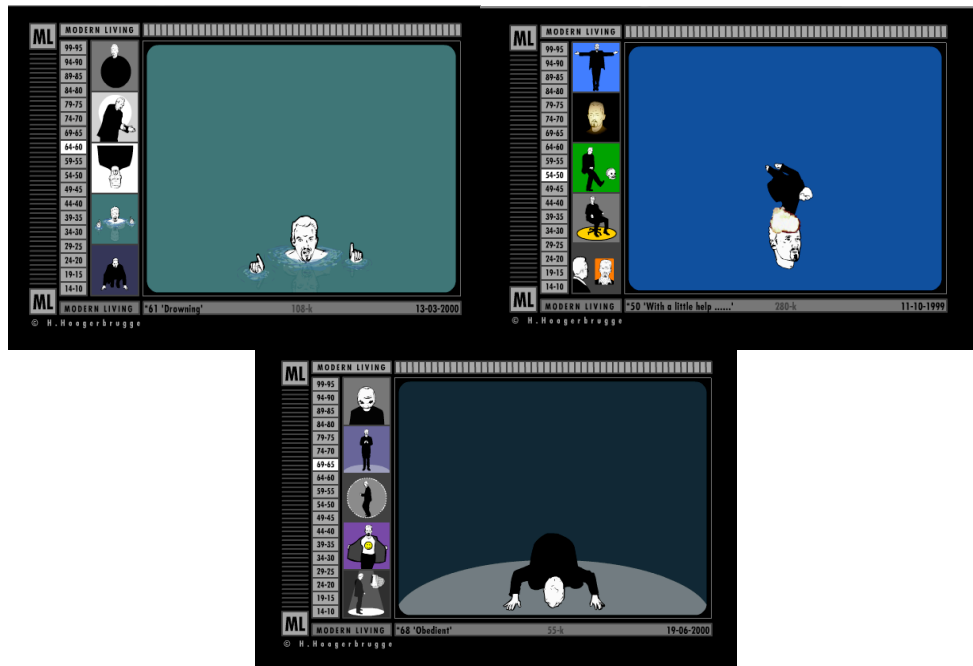


Figure 6.24 #61 'Drowning', #54 'Jumpy', & #68 'Obedience' in *Modern Living*

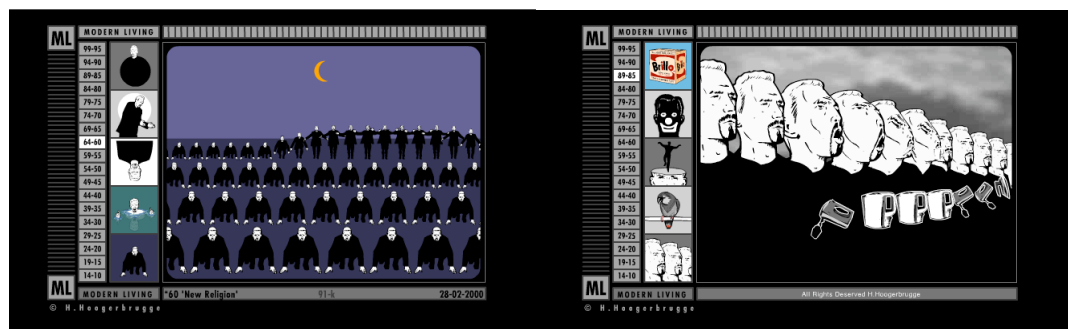


Figure 6.25 #60 'New Religion' & #85 'Material Guy' in *Modern Living*

Understanding

As mentioned, many pieces in *Modern Living* remap conventional mouse actions to novel meanings. These new mappings are formed with the help of immediate blending. Furthermore, some pieces even provoke metaphorical understandings. To illustrate, I focus on an example. In #83 'Possessed', the character sits behind a task table facing the user. It seems like he is using a mouse. When the user puts the pointer on the task table, the character would move his mouse to follow. The immediate blend here lets the user easily control the character by directing him to desired location, like the mechanism of

controlling shadow puppets by moving rods. The analogy link between the character and a rod puppet is compressed into a concept of virtual puppet in the animated piece. Hence, when one moves the pointer, the virtual puppet moves his hand.



Figure 6.26 #83 'Possessed' in Modern Living

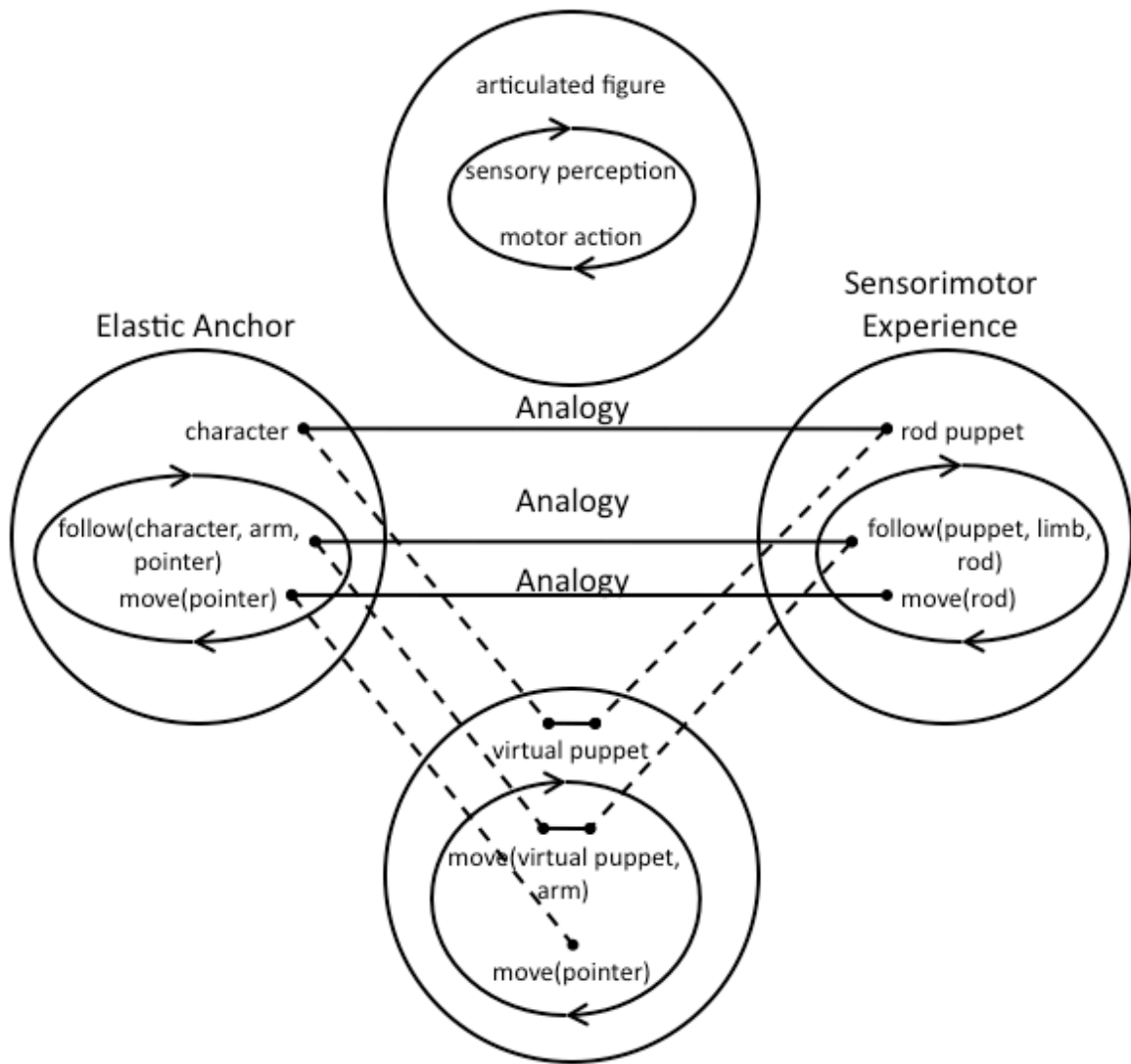


Figure 6.27 The immediate blend in #83 'Possessed'

What makes the piece eccentric is that the character keeps popping bizarre appearances over his head and making curious sounds while the user moves the pointer around. This effect invokes the “web surfing” frame in the user’s mind, suggesting an metaphorical blend of the animation with an imaginative story in which people surfing the Web, including the user, often have their attentions wandered and diverted by many hyperlinks. In the blend, the mapping between the character and the user, both are moving a mouse, is compressed into an inner-space identity relation. The animation becomes like a mirror reflecting the mental status of the user. This example shows a typical way of how *Modern Living* echoes our personal experiences.

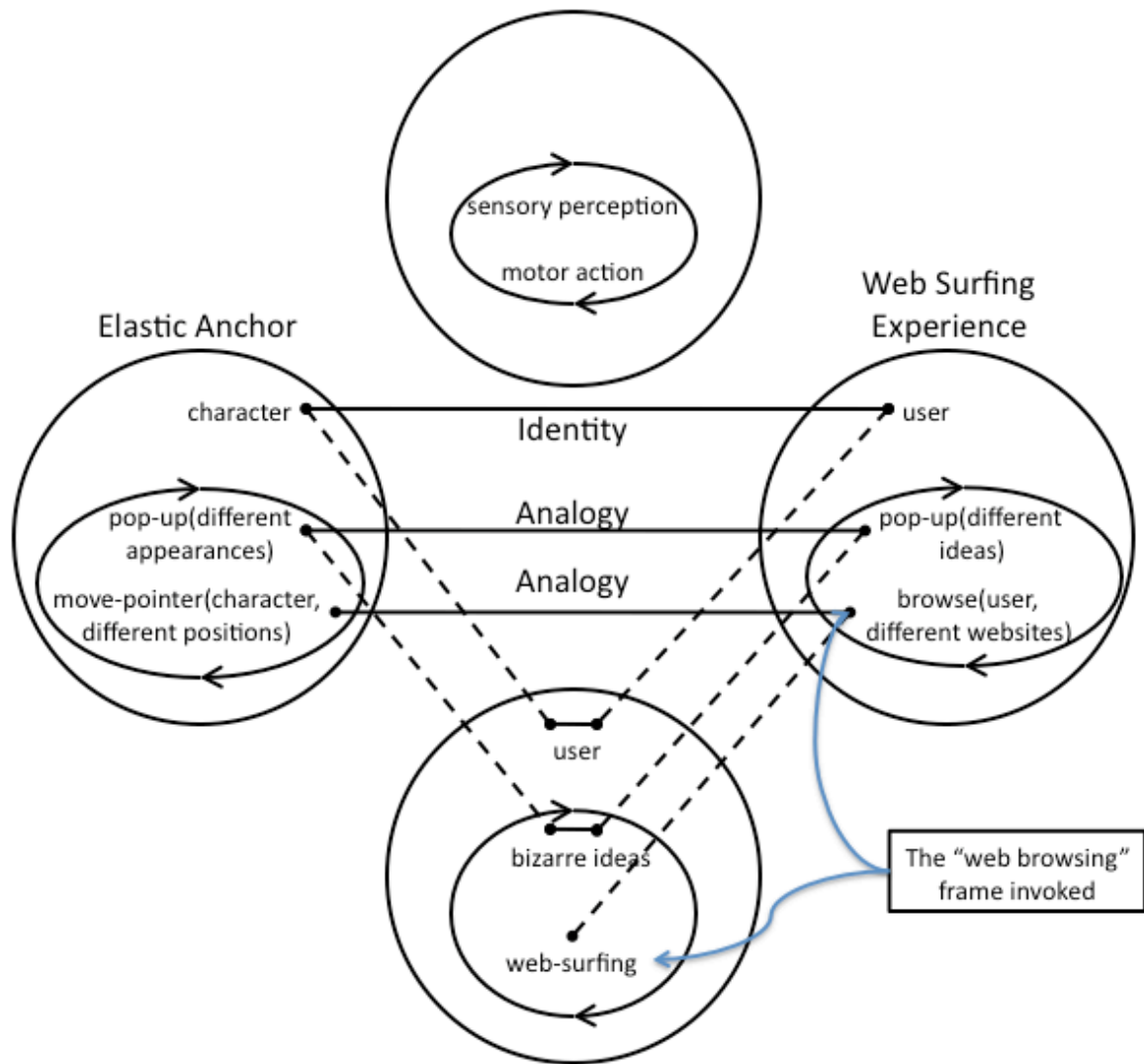


Figure 6.28 The metaphorical blend in #83 'Possessed'

Bodily Engagement

In *Modern Living*, most animated pieces are interactive. The system considers the position of the pointer and the timing of a click in user input. For example, in #70 'Eternal Love', the position of the pointer determines whether the character in the foreground or the picture in the background is in focus. In #87 'Vaudeville', many "dummy" versions of the character travel across the frame at various speeds. A user can point and click on any heads to see the pop-up, like playing a classic carnival game.

Whether a head pops certainly hinges on the motion and timing of the user's motor action. These interaction mechanisms can be considered slightly motion-sensitive.



Figure 6.29 #87 'Vaudeville' in *Modern Living*

In case of no user action, most pieces just come to a halt. Even though a few pieces do show some changes, like #63 'Perfect Day' in which the character keeps swinging along the background music, the audio and the visual are just in very simple loops. The environment does not really transform into next stage until the user moves the pointer. It is not an enduring environment. Hence, *Modern Living* only entails decelerated engagement.

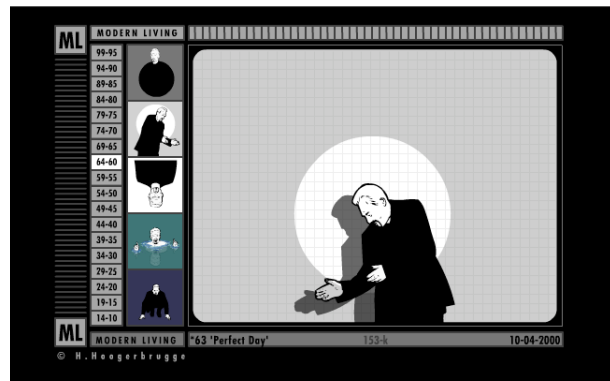


Figure 6.30 #63 'Perfect Day' in *Modern Living*

Liveness

Although *Modern Living* is regarded as interactive web comics, it also opens up possibilities of improvisation between the user and the artist. An example piece best

demonstrating this idea is #71 'El Mariachi'. The piece features on stage five characters, each of which holds a guitar. A user can cast a spotlight on any particular character, and that character would start to play. Although all characters and guitars look similar, each of them makes a different sound. Hence, by moving the pointer the user is like playing a synthesizer with five different instrument tracks. The result is a performance. In this performance, the stored components include the audio and the animated drawings prepared by the artist. Meanwhile, the spotlight effect is still added in real time. The performance can be said as slightly live.

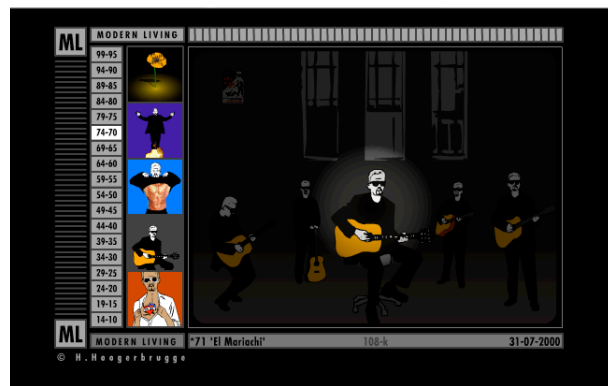


Figure 6.31 #71 'El Mariachi' in *Modern Living*

This section has talked about two exceptional digital artworks. They differ from each other in presentation format, intended message, and representational style. I include both of them because they both rely on generative and interactive animation to project an expanded illusion of life, as the above analyses has proved. Since the two works are also widely covered in the literature of digital art, they are very influential in their respective sub-domain. In this regard, I believe there are many more followers, and they altogether continue defining the new animation phenomena in the digital age.

Conclusion

This chapter has conducted close reading on a wide array of media artifacts, which even come from such basic-level media categories as animation films, video games, entertainment applications, user interfaces, multimedia websites, digital signs, and

digital artworks. The survey mainly includes descriptions of the artifacts and analyses of their properties with respect to the four principles of generative and interactive animation. The results can be summarized in terms of the four variables and juxtaposed for comparison in the following table:

Table 6.1 Variable values taken by the corpus

Basic-level Category	Artifact	Type of Liveliness	Level of Understanding	Mechanism of Engagement	Degree of Liveness
Short animation films	Arabesque	emergent	immediate, metaphorical	n/a	playback
	Luxo Jr.	interactional, intentional, emergent	immediate	n/a	playback
Video games	Pong	interactional, intentional	immediate	decelerated	highly live
	Passage	interactional, intentional	immediate, metaphorical	continuing	fairly live
Entertainment applications	Koi Pond	interactional, emergent	immediate	moderately decelerated	fairly live
User interfaces	Genie effect on Mac OS X	intentional	immediate	halted	live
	Dock on Mac OS X	intentional, emergent	immediate	decelerated	live
	Water-level interface on N702iS	emergent	immediate, metaphorical	continuing	highly live
Multimedia websites	SnowDays	emergent	immediate, metaphorical	continuing	live
	Come Clean	intentional	immediate, metaphorical	halted	fairly live

Digital signs	Walking green man on traffic lights	intentional, emergent	immediate	n/a	slightly live
Digital artworks	Text Rain	interactional, intentional, emergent	immediate, metaphorical	continuing	highly live
	Modern Living	interactional, emergent	immediate, metaphorical	decelerated	slightly live

Based on the above table, we can examine the distribution of the variable values among the corpus of artifacts. There are three observations.

First, the four variables are orthogonal. No one variable seems to take a particular value tied with another of any other variables. For example, consider first the type of liveliness and focus on those artifacts showing emergent liveliness. They might provoke just immediate understanding (e.g., *Koi Pond*), or multiple levels of understanding (e.g., *SnowDays*). Some of them entail decelerated engagement (e.g., *Modern Living*) while some others continuing (e.g., the water-level interface). Their respective degrees of liveness can be highly live (e.g., *Text Rain*), slightly live (e.g., animated traffic lights), or even purely playback (e.g., *Arabesque*). Locking down any one variable also would not render us any obvious pattern in other three variables. Hence, these variables are independent of one another as far as the corpus is concerned. In short, the results demonstrate my initial suggestion of looking at the new animation phenomena from the four orthogonal perspectives and latter theoretical development of each principle.

Second, no one basic-level category has bias toward certain value in any particular variable. In contrast, each category has spanned a fairly broad terrain in all the variables. For instance, most categories have some artifacts involving continuing engagement, as well as some others decelerated. Even for a category like entertainment applications that only contains one artifact, the single sample also shows both

interactional and emergent types of liveliness. On the one hand, it follows that although now the sample size of each category is kept minimum in the corpus, it is still able to encompass the largest possible variety in each quality of the new animation phenomena. On the other hand, this finding tells us that each category has equal potentiality in demonstrating an expanded illusion of life.

Finally, with all samples considered together, each variable has taken all its possible values without apparent prejudice. For instance, about half of the samples provoke both immediate and metaphorical understandings while the other half only immediate. Meanwhile, nearly one third of the samples entail continuing engagement and another one third only decelerated. Although in Chapter 4 I painstakingly advocate the four principles, including material-based imagination, enduring interaction, and so on, the unbiased presentation of values in this chapter show that it is no easy task to have an artifact exemplifying in all four variables, that is to say, multiple levels of understanding, continuing engagement, and so forth. Sometimes, an artifact might be able to provoke multiple understandings but only involve halted engagement (e.g., *Come Clean*), or an artifact with continuing engagement might not demonstrate all types of liveliness (e.g., *Passage*). Therefore, the samples tend to compensate one another, resulting in a featureless distribution. All in all, this observation implies that the four principles still set a challenging and ambitious objective for designers and creators of digital media artifacts to achieve in the future.

The text and figures of this chapter, in part, contain material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2009a, 2009b). The dissertation author was a co-author of this paper.

CHAPTER 7

DESIGN IMPLICATIONS

In previous chapters, I introduced new principles of animation each of which is articulated by a qualitative variable. I then applied these variables with different possible values to a corpus of media artifacts, including typical and salient examples from various basic-level media categories. The attributes of these artifacts conceptually spread and cover a terrain across the multidimensional space in which each dimension denotes one variable and the corresponding possible values distribute along the axis. The terrain reflects the inclinations of the new animation paradigm exemplified by the corpus.

Most important of all, the absence of any bias in variable values lets us know that not every existing animated artifact can epitomize all four principles of generative and interactive animation, and so there is still room for multimedia designers and developers to explore and improve. Hence, for designers and developers, the principles and variables work as an index informing the position of a particular work with respect to the whole existing terrain. Based on this relative positioning, designers or developers can set adaptive and situational benchmarks for evaluation and assessment. Meanwhile, the multidimensional space not yet covered by the terrain can remind them of new possibilities for more embodied, evocative, and affective forms of interactive generative animation. People can determine their ambitious and feasible goals with respect to constraints in resources and time.

Practically, the set of principles and variables suggest new design possibilities (perhaps even paradigms) in two areas. On the one hand, some parameters like emergent liveliness, metaphorical understanding, and so on suggest provocative approaches to designing animation and motion graphics in addition to the classic principles of animation introduced by Disney. On the other hand, such new concepts as elastic

anchors, motion-based motor input, puppeteerly animation, and so forth, constitute a new set of interface design principles rooted in animacy, proposing an alternative to the prevalent user-centered and the contending agent-based approaches. This chapter describes these two design implications.

Motion and Animation Design

Today we can see animated images on many different types of screen, such as animated films in cinemas, motion graphics on television, computer graphics on video games, Flash movies on the Web, commercials on electronic billboards, and the like. In contrast to these diverse application domains, the major principles of animation most animators follow are still limited to the Disney's heritage in the cel-animation era. However, if one looks at typical works of Japanese animation, it is not difficult to notice that most Disney's principles do not apply, as will be shown below. The disparities are even more obvious when animation is juxtaposed with motion graphics, as we have seen in many movie titles, music videos, commercials, outdoor projections, and so on. In this regard, the principles of generative and interactive animation are able to accommodate the varieties of general motion and animation design in the digital age.

Disney's Principles of Animation and the Rival

Based on over decades of experiences, Walt Disney Studios established the well-known twelve principles of animation. Although emerging from the cel-animation context, the principles are so well recognized that they are regarded as dogma even in computer animation production. Yet, the principles almost only address character animation. Out of the twelve principles, many, including Anticipation, Staging, Exaggeration, and Appeal, aim to make action or gesture clear to audiences. As the noted Disney animators Frank Thomas and Ollie Johnson put it, Walt Disney called this approach "aiming," which is one of the first Walt's insistences (Thomas & Johnston,

1984, p. 20). In my words, it corresponds to intentional liveliness. Meanwhile, a few other principles, like Squash and Stretch, prescribe the conservation of volume in an animated body while changing shape in order to show elasticity and register weight (Thomas & Johnston, 1984, pp. 18,30). It is partly concerned with interactional liveliness. These principles concentrate audiences' attentions on prominent postures or actions. The few others pertinent to diluted attention are Follow Through and Secondary Action. The former states that the loose flesh on a moving figure would move slower than the skeleton, while the latter suggests adding subsidiary actions to the dominant action. These extra details help distribute certain audiences' attentions to scattered parts of the main figure, adding "richness to the scene" (Thomas & Johnston, 1984, p. 32). Unfortunately, the Disney animators also unintentionally disparage this kind of "effect animation" as "mechanical" (Thomas & Johnston, 1984, p. 29).

Conversely, some rivals of Disney celebrated the mechanical, such as the one-time contender Fleischer Studios. In order to compete with Disney in the beginning of the battle around the 1930s, the Fleischer animators dismissed the Disney style and attempted their own way. For instance, in *Betty Boop's Snow White* (1933), we cannot easily find an example of Anticipation or Squash and Stretch that follows the laws of physics, but instead we see many rhythmic animation cycles and metamorphoses. Norman Klein has repeatedly juxtaposed Disney and Fleischer in terms of metamorphosis and the human-machine opposition (Klein, 1993, 2004). He sees Disney as flesh and Fleischer automata (Klein, 1993, p. 76). The former insists an illusion of life that hides the animator (Klein, 2004, p. 254), while the latter celebrates technologies of animation and self-reflexivity (Klein, 1993, p. 168, 2004, p. 255). Regarding shape-shifting, Disney required animators to keep a character's volume constant, whereas the Fleischer animators were free to transform their characters' bodies into any form. For example, when the clown in *Betty Boop's Snow White* is turned by the evil queen into a ghost, he involuntarily keeps dancing in "loops" and morphing without bound, even into a medal in chain. Klein

describes this type of morphing phenomena as hesitation or lapse, for it is always unstable and seemingly reversible (Klein, 2004, p. 254). These seemingly reversible phenomena demonstrate emergent liveliness, and Klein's use of the term "lapse," which literally means temporarily loss of concentration, particularly resonates with my idea of diluted attention. Hence, Fleischer is more obsessed with emergent patterns and acts like a perfect complement to Disney's emphasis on prominent behaviors, forming a holistic illusion of life.

Japanese Anime

Some might argue that the failure of Fleischer Studios and the prevalence of the Disney style suggest that emergent liveliness, which dilutes our attentions, does not appeal to mainstream audiences. This argument is revealed to lack grounding when we switch the context to another side of the globe. Japanese animation, commonly called "anime," represents a substantial part of Japanese cinema that not only has established massive markets in Asia but also has started to gain increasing international recognition. When one looks at an anime, regardless of a television series of limited animation (animation done in less than 24 frames per second, usually 12), a Hayao Miyazaki's feature film, or an OVA (original video animation released directly to home-video formats), it is easy to notice the striking differences in visual composition and animation style from the western counterparts. Generally, anime works might have more and longer holds of characters without movement, more animation loops, more planar movements of multiple layers composited together to imitate camera movement (based on a visual phenomenon called motion parallax) (Lamarre, 2009, p. 39), and less continuous and complete character action without cut. Part of these phenomena, like loops and holds, are believed to be results of the low budgets facing most anime productions in the post-WWII period, arguably set by the successful precedent case of Osamu Tezuka's *Astro*

Boy released in 1963, and later turned into a distinctive and well-received style of anime (Ono, 2007, pp. 106-107).

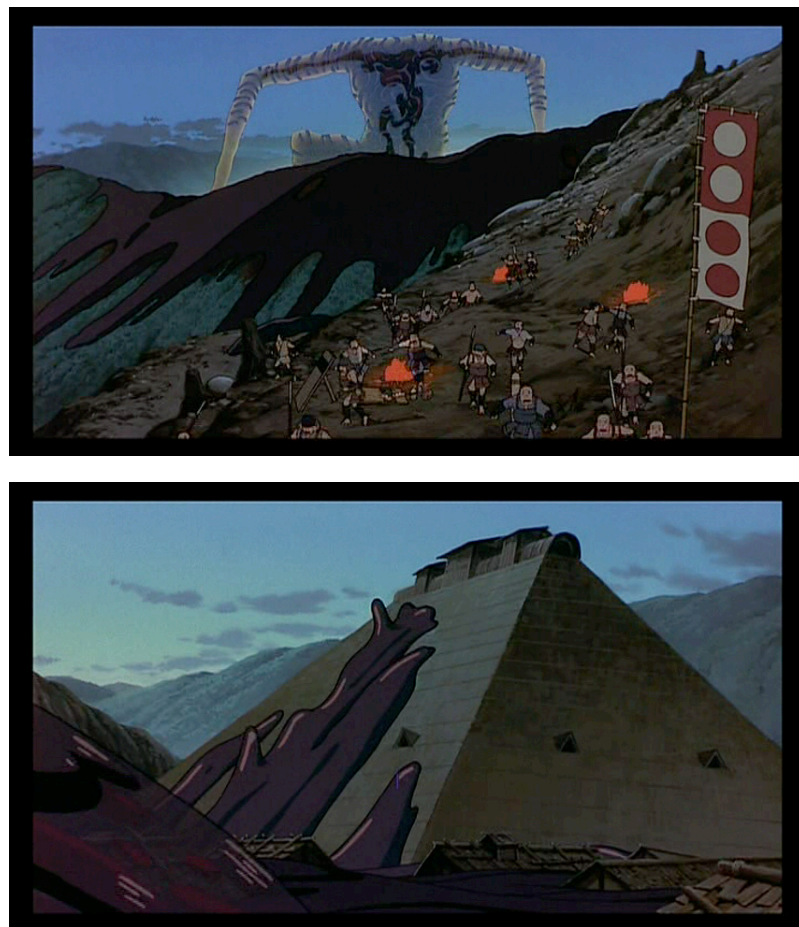
Interestingly, this signature of anime manifests the vitality of emergent liveliness in animation phenomena, especially when loops, holds, and parallax are concerned. An example would help illustrate the point. Consider a three-minute sequence (Figure 7.1) in the Mamoru Oshii's animated cyberpunk film *Ghost in the Shell* (1995). The whole three-minute features primarily emergent liveliness, diluting audience attention over different parts of the frame rather than on particular figures. The sequence exhibits nostalgic images of a city with many recognizable Hong Kong elements, like flyovers, narrow streets, commercial billboards, dashing cars, and trotting crowds, plus imaginary tram-boats and canals. In many shots, it applies multilayer movement, like steady flow of canal water with reflection of a moving boat, billboards running across in the foreground while people working in a commercial building at the background, pedestrians passing by in front of a shop display reflecting some animated commercials, and so forth. Dynamics are distributed over the frame, and movements seem reversible with meaning intact, constituting typical emergent liveliness. Although the protagonist appears in some shots, she stays relatively still, with moving billboards at the background, her hair swaying in the breeze, or varying window reflection in the front. In the later half of the sequence, raining adds even more emergent liveliness to the seemingly timeless city through lines, ripples and running highlights (Figure 7.1). All in all, the sequence exemplifies emphasis of anime on lively atmosphere over progressive action.



Figure 7.1 Screenshots of *Ghost in the Shell*

Disney's emphasis on character animation, Fleischer's love of morphing and technologies, and the distinctive style of anime represent different inclinations toward different types of liveliness. Yet, these approaches to animation are by no means clearly separate from one another, but actually there are rising mutual influences. When seeing the computer-generated character Sulley in Pixar's *Monsters, Inc.* (2001) acting

prominently with all his three million hairs moving individually, or the Forrest Spirit-turned headless gigantic body in Miyazaki's *Princess Mononoke* (1997) looking for its head with deadly fluid extending everywhere in the landscape (Figure 7.2), one would be aware that today's animation tends to manifest a holistic sense of liveliness. Hence, in designing animation, even character animation, the new principle about holistic illusion of life reminds designers and animators of the vitality of the usually denigrated emergent atmosphere in balanced relation to the prominent action celebrated by the classical principles.



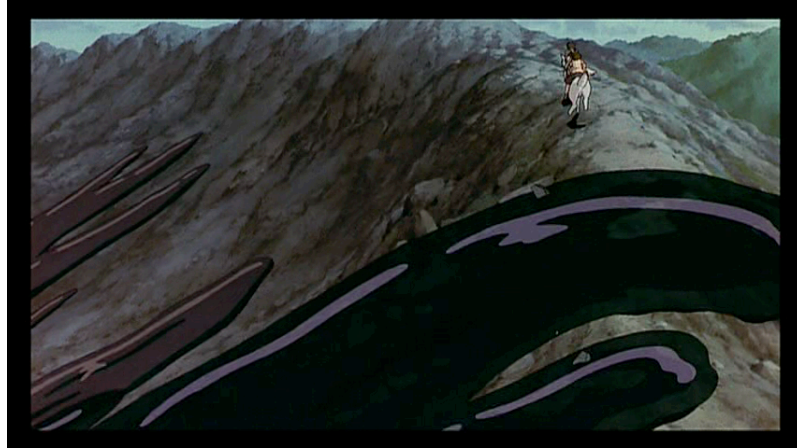


Figure 7.2 Screenshots of *Princess Mononoke*

Motion Graphics

This holistic sense is also commonplace in a cousin of animation, motion graphics, with its increasing reliance on computing technologies in creating an illusion of motion. The term arguably emerged since John Whitney started to produce abstract geometric animation by his computers and the graphic designer Saul Bass applied animated graphics to create title sequences for movies. Their most notable collaborative work is the motion title of Alfred Hitchcock's *Vertigo* (1958), which features swirling patterns superimposed in a woman's eye, symbolizing an unstable state of the mind. The uses of emergent patterns in moving images, on the one hand, add diluted attention to concentrated attention, forming a holistic sense of liveliness, on the other, help provoke multiple levels of understanding the images. The ideas can be better illustrated by Bass's another outstanding work, the title of Hitchcock's *North by Northwest* (1959). The sequence starts with a green background of grid lines. The film title and names of featured actors consecutively move up and down with several-second holds in the middle of the frame. A prominent vertical movement of a name is often accompanied with a horizontal paddle moving in the adjacent column. Soon after, the grid dissolves into the mirrored façade of a building, reflecting busy yellow cabs traveling horizontally on the grid. These additional animated graphic elements not only bring in emergent liveliness to

balance concentrated attention, but also provoke metaphorical understanding that the moving types are elevators and the paddles are the balancing weight. Other following shots, like flocking crowds in the streets and finally the main actor walking out of an elevator, help invoke the frame of the story: fast-paced chasing in an urban city.



Figure 7.3 Screenshots in the title sequence of *North by Northwest*

Whitney, Bass, and other designers like Maurice Binder, the creator of many James Bond movie titles, have paved the way for combining motion design intriguingly with moving images as title sequences. Unlike feature animation, these sequences usually do not recount narratives in the narrow sense, because there is no obvious action and reaction, nor cause and consequence being unfolded. Instead, they convey meanings more through emergent patterns and metaphorical understanding. Consider one of the latest James Bond movies, *Casino Royale* (2006). In the film title, we see gorgeous moving patterns of poker suits, intertwining with metaphorical transitions, including a Queen to a lady, a Spade to a bullet, a Diamond to a killing dart, a roulette wheel to a riflescope, and

others. These emergent motion graphics invoke the dark side of the special business, building a background atmosphere for the audience to go into. Today motion graphics have been spreading widely to television, the Web, music videos, commercials, and even art and literature⁹. A major difference in context from animation is about duration. Motion graphics usually consume just very limited time, but convey succinct meanings and project emotional ambience. In this regard, holistic consideration of liveliness and metaphorical understanding advocated in the new principles would help motion designers synthesize and evaluate their creation with strong theoretical underpinning.

Interface and Interaction Design

Apart from animation design, the new principles, especially the latter two about the body and the performer, also give us new implication on interface design. In the analyses of the corpus of media artifacts, I show that generative and interactive animation opens up possibilities of creating interfaces that engage users in a lively, imagination-provoking, continuing, and contingent environment. I propose that this kind of interfaces can throw new light on the two contending approaches to HCI, namely direct manipulation interface and intelligent software agent.

Direct Manipulation Interfaces

In interface design, one of the primary goals is to make interfaces more “user-friendly,” or “usable.” Following the views of Donald Norman (Norman, 1988), Jakob Nielsen (Nielsen, 2000), Ben Shneiderman (Shneiderman & Maes, 1997), and others, the terms generally describe a tool that is easy and rapid to learn, understand, operate, or

⁹ For example, the Electronic Literature Organization published a collection of works in 2006, many of which include motion typography as an integral presentation element, such as Brian Kim Stefans’s *The Dreamlife of Letters*.

control. Their ideas have led to the prevalent user-centered approach to interface design. The paradigm emphasizes usability, efficiency, and productivity.

I take the perspective that user-friendliness can be interpreted with the broader perspective of user-centricity. The quality depends on which type of users the system is targeting to. To expert users, granting full access to the system projects a sense of full empowerment. The user feels that everything is in control and for his or her “direct manipulation” (Shneiderman, 2003 [1983]). For example, in the MS-DOS command-line environment, an experienced user is able to directly “order” the system to take a particular action and might think that the system is straightforward and efficient to use. This kind of “manipulation” is in a metaphorical sense. In contrast, this empowerment might overwhelm a novice user, who is not familiar with the system and not used to directly controlling it. To help this type of users, Norman’s idea of natural mapping comes into play (Norman, 2002, pp. 75-79). By mapping an intuitive, ordinary, and everyday environment to the system, the interface becomes familiar to the novice user. The most illustrative example is the GUI environment, which “imitates” the physical working environment consisting of “desktop,” “folder,” “trash,” and the like. This “imitation” hinges on the user’s mental operation, Fauconnier and Turner’s “conceptual blending,” which is largely preconscious, mapping various interface actions to everyday physical actions, such as “click and drag” to “grasp” (Fauconnier, 2001). Through this immediate blending, the user feels directly manipulating the components in the system as natural as what one usually doing in the physical environment. This is a literal sense of “manipulation.” The two senses may correspond to different design solutions, but both of them are also user centric, seeing machines as instruments, aiming to build systems easy and efficient to use.

Intelligent Software Agents

Whether targeting the expert or the novice, designers of user-centric interfaces regard computers as an instrument for users' direct manipulation. Another contending approach to "user-friendliness" is to make it really like a friend, a companion you are working with, or even a delegate working on your behalf. The former major proponent Pattie Maes believes that today's computer users need this kind of software delegates to give them advices because most of them are novices and they are overloaded with information in digital environments (Shneiderman & Maes, 1997). Yet, this agent-based approach is relatively ambitious, if not futuristic, because being a qualified representative, the agent has to be intelligent enough to make prompt decisions favorable to its master. It is about not only designing an interface, but rather designing an "intelligent" system. This idea emerged within artificial intelligence research, often using symbolic AI approaches, over the past two decades. Earlier examples of seemingly intelligent software agent include Joseph Weizenbaum's conversational program, ELIZA. Yet, the program, rather than serving as a delegate, is an attempt to cheat its users into believing that "she" is a human psychotherapist, in other words, an address to the famous Turing's test. Later, other projects started to import AI techniques to create "smart" software assistants, such as ALICE the chatterbot (Suchman, 2007, p. 216) and not to mention many of Maes's works. This stream has triggered once charged interface versus agent debate in the field.

In fact, since the Turing test, to latest critiques by Hubert Dreyfus (Dreyfus, 1972), critics have informed that strong artificial intelligence, which aims at general human intelligence, is a challenging task, because the human mind is not just a pre-programmed information-processing unit but actually incorporates other situated and embodied cognitive processes. Implementing intelligence with a software model would just limits and compromises a user's behaviors to that model. Jaron Lanier even cynically stated that such a "smart" system is only possible if we, the human user, "make ourselves dumb!" (Lanier, 1995) To this, Maes contends that sometimes users just want to leave

problems to experts (and to be ignorant), and so they can do other things instead (Shneiderman & Maes, 1997). Hence, agents would let every user focus on another levels of interest or expertise.

However, the idea of intelligent agents might not be welcomed by all users, as reflected in many western science fictions and movies, like the cyborg monster in Mary Shelley's *Frankenstein*, or the Hal computer in *2001: The Space Odyssey* (1968). As Sherry Turkle puts it, if the computer is too smart, too humanlike, it becomes unpredictable. This "psychological machine" might not be a friend, a safe companion, or a faithful delegate of humans (Turkle, 1984, p. 238). The cyber-culture writer Steven Johnson also pictures the overwhelming and chaotic future resulted from the abuse of agent-based technologies by so-called "push media" and the turbulence caused by the notoriously unpredictable feedback mechanisms (Johnson, 1997).

An Animacy-Oriented Approach

At the level of vision, the interface approach differs fundamentally from the agent-based approach in that the former treats computers like a tool or an instrument while the latter regards computers as an assistant. At the level of application, both of them are in fact inadvertently concerned with productivity or usability of a system, making computer a synonym of office equipment or business machine. They share similar application domains, primarily utilities.

However, with latest advances in multimedia and information technologies, the meaning of application is not limited to utilities. Today we have numerous software applications for amusement, education, social communication, and even artistic expression, as I have illustrated in the corpus of media artifacts (see Chapter 6). People dealing with these applications are not just users, but also players, participants, performers, and "interactors." The utilitarian criteria like productivity, efficiency, rate of errors, and so on might not fully apply to these types of computer systems. Hence, the set

of principles of generative and interactive animation I have described suggests an alternative approach to designing these emergent paradigms of digital media artifacts. The principles about animacy and imagination highlight how liveliness and motion in both motor input and sensory feedback allow digital media artifacts to convey non-verbal meaning to users through elastic anchors. For example, considering the types of liveliness lets designers of a system know if the animated feedback concentrates a viewer's attention on the progressive action signaling a change of state, or dilutes one's attention in an imaginative story world provoked by the system. Furthermore, the level of understanding tells designers how provocative the tie between the motion-based motor input and the animated feedback is, and how it constitutes an elastic anchor for imaginative blending.

Meanwhile, the notion of enduring transformation in digital environments emphasizes the use of continuing engagement to give users a sense of spatiotemporal embodiment in the virtual world. For instance, by reviewing the mechanism of bodily engagement, designers can assess which specific parts of motion or animation in the interface contribute to providing embodied experience of an illusion of autonomously and continuously changing life. Finally, liveness reminds designers of the contingency of a system. A highly live system means that it is more open to user intervention, more adaptive and personalized, even to the extent that Maes supposes software agents to be (Shneiderman & Maes, 1997).

Some might think that the principles I have described are just another attempt to anthropomorphize the computer. Shneiderman particularly dismisses the idea of adding animation to user interfaces. He thinks that “anthropomorphic representations destroy the users' sense of accomplishment,” making them feel like lost control of the systems (Shneiderman & Maes, 1997). To Shneiderman, and even Maes, the term “anthropomorphic” implies imitation of all sorts of human activities including thoughtful speech and intentional act. Making machines performing these activities are ambitious

goals and might lead to undesirable outcomes. Therefore, my suggestion is that we can aim at a kind of “softer” anthropomorphism in animacy-oriented interfaces, which imitate not human intelligence but just liveliness of animate beings and divergent transformation of environments. What I mean here is to make human-computer interaction more like humans’ everyday experiences with life. This kind of anthropomorphism represents a middle ground between seeing the machine as an instrument or turning it into an intelligence. The aim should not be to make machines act “as” humans, but rather to act “familiar” to humans.

As I have mentioned in Chapter 4, Joseph Bates and his team have consciously related Disney’s principles of animation to their aim of creating believable characters (Bates, 1994). They proposed that emotion is one of the significant qualities in this make-believe process, so the team put their emphases on how to represent emotions internally and develop AI architecture for generating and expressing emotions. In fact, it is just another symbolic AI-oriented mission. As Suchman puts it, “Bates’s creatures are simultaneously presented as illusions of life *and* as important steps along the path to the real thing” (Suchman, 2007, p. 212, original emphasis). I add that it is more about a simulation than an illusion, more on aliveness than liveliness. In contrast, my proposal suggests a shift of focus toward perception of animacy and embodied cognition on the human side in the interaction mechanism (though computer representation and generation still persist on the other side). I believe that interface designers can eschew the goal of designing seemingly intelligent systems while still taking advantage of liveliness and embodied understanding enabled by animated elements in an enduring environment to project an image of computers that is more familiar, intimate, and close to humans, that is, a general sense of anthropomorphism, which is softer in approach but strong in vision.

Conclusion

This chapter has looked into two inherently separate design disciplines, in which animation design is generally related to the film art while interface design primarily emerges from the field of HCI. However, as I have demonstrated, with today's accessible and ubiquitous computing technologies, interactive generative animation connects the two lineages of design, expanding the theoretical perspective of animation design and bringing promising approaches to interaction design. To designers, this interdisciplinary study of animation and interaction recentralizes humanistic theories and interpretation in the analysis and design of multimedia artifacts, giving rise to more "human-familiar" and "life-like" design solutions. To developers, the integration of computational and cognitive research results with interpretive approaches from humanities disciplines provides a new orientation for designing technologies that are more in line with our everyday and intimate experiences.

CHAPTER 8

THE GENERATIVE VISUAL RENKU PROJECT

Previous chapters covered the theoretical aspects, analyses, and implications of interactive generative animation. The new design principles advocate presenting a holistic illusion of life including both prominent behaviors and emergent patterns, provoking multiple levels of imaginative understanding, engaging users in motion-based input and enduring environments, and providing them with divergent and contingent experiences. To illustrate these ideas, this chapter describes an ongoing digital artwork project related to the new animation paradigm. It starts with introduction of the theoretical and artistic background, followed by discussions on the design considerations in generative algorithms, transformational rules, and interaction mechanisms with respect to the principles. Lastly, the corresponding variables will be articulated with design evaluations. In short, this chapter demonstrates how the design framework resulted from previous chapters can be applied to creating generative and interactive animation phenomena in digital artworks, and also multimedia artifacts at large.

The Generative Visual Renku (GVR) project, developed by D. Fox Harrell and the author, presents a new genre of generative and interactive visual art inspired by generative models from contemporary art, Japanese renku poetry, and an account of the interplay between iconicity and conceptual metaphor by Masako Hiraga and C. S. Peirce's semiotics (Hiraga, 2005). Works of this new genre generate expressively meaningful compositions of images dynamically through interactive co-creation between the user and the system. Generative Visual Renku, as a genre, could be considered an example of a digital media practice that Harrell terms "Phantasmal Media," wherein imagery (broadly construed) and ideology intersect in computational processes of meaning construction and transformation (Harrell, 2009).

Traditional renku is a type of linked haiku poetry, consisting of a series of links between topical elements in verses. The visual renku described here uses animated images in place of traditional written text, and the linked disposition of these images is maintained algorithmically. It is generative in that algorithms are used to dynamically compose variable instances of imagery. It is also interactive, outputting different instances of imagery in response to different user actions upon depicted visual elements. The current GVR implementation is built upon the GRIOT platform, a computational system designed by Harrell to support semantics-based interaction and generation for multimedia works (Harrell, 2007). Initial cases of GRIOT application, however, were primarily textual. One of the objectives of the GVR project is to extend the GRIOT architecture to implement meaningful constraints for visual composition (Harrell & Chow, 2008). Meanwhile, the project aims to explore possibilities of the computational framework based on cognitive semantics accounts of imagination engaging the cultural practices in which meanings in both visual and mental imagery are conveyed in some non-Western discourse artifacts, such as Chinese character forms and Japanese renku poetry.

As I have mentioned in Chapter 2, generativity and interactivity jointly give rise to phenomena of liveliness, including autonomous transformations and reactive behaviors. During the GVR project development, we found that the new genre inevitably entails these new animation phenomena (Chow & Harrell, 2009a) and so, designing GVR works also demand considerations of the new principles I have proposed. This chapter first introduces the artistic and theoretical framework of the project, and then focuses on an example of generative visual renku, followed by discussions on the application of the new principles.

Artistic and Theoretical Framework

The GVR project draws upon an interdisciplinary theoretical framework including generative visual and literary arts, Japanese renku poetry, cognitive science and computer science research results, and Chinese character forms. In (Harrell & Chow, 2009), Harrell and the author comprehensively delineate this framework. The following is a brief summary.

Generative Visual and Literary Arts

Generative art or literature generally refers to a type of creative works that strongly emphasizes the processes giving rise to specific instances. The process is usually designed by the artist or writer in terms of a set of rules, a computer program, or a machine, and then is executed automatically (Galanter 2006) (For more discussions on Galanter's definition of generative art, please see Chapter 2 and Chapter 4). The Oulipo (Workshop for Potential Literature) group is perhaps the most prominent proponent of this concept in literature (Mathews & Brotchie, 1998). In the case of visual art, many contemporary art practitioners have created inspirational exemplary works, such as Josef Albers' *Homage to the Square* (the 1960s), Frank Stella's *The Protractor Series* (1967-69), Sol LeWitt's *Incomplete Open Cubes* (1973-4), Jennifer Bartlett's *Rhapsody* (1975), and Xu Bing's *Book from the Sky* (1988). In such works, the aesthetic sensibility focuses on directing viewers to considering processes rather than individual instances of output. Meaning is embedded in the form as opposed to content. Although opponents criticize that these formal avant-garde approaches to the arts are often emotionally dry, we found that the related generative models, such as the procedural, serial, and combinatorial methods, are excellent references for the design of generative rules in the GVR project.

The Link and Shift in Renku

Linked poetry is a kind of collaborative literary work. It requires two or more poets to alternate in the creation of verses to form a complete poem. In Japan, renku, a

form of haiku with linked verses in the style of the Basho school, is highly praised as a national heritage and is growing in popularity in international literary circles (Higginson 1996). To compose a poem, renku poets have to create “links” between successive verses through object, meaning, or scent, which mean a consistency in mood or emotion. Meanwhile, “shifts” are used to avoid throwbacks or too many repetitions of ideas or themes. The key is to balance connection and diversity. The poet’s choice of words then depends not only on personal aesthetic sense but also the context.

A classic example of renku poetry by Matsuo Basho follows (with added bracketed words to indicate linking and shifting) (Higginson, 1996):

Around the town
the smells of things ...
summer moon
[link]
“It’s hot, it’s hot” -
the voices from gate to gate
[shift]
the second weeding
not yet done, and ears
out on the rice

The first and the second stanzas are linked by the summer season and its hotness. The second stanza is connected with the third stanza through the auditory – voices to ears, but the location is shifted from the residential area implied by the multiple “gates” to the rice field.

Cognitive Semantics: Metaphor, Blending, and Analogy

George Lakoff, Mark Johnson, and others have studied metaphors as mappings between domains and have shown that many basic metaphors are based on everyday life

experience (Lakoff & Johnson, 2003). Gilles Fauconnier and Mark Turner have extended this framework in their conceptual blending theory, which describes how concepts are integrated both unconsciously in everyday life and in more complex thought such as in art and literature (Fauconnier & Turner, 2002; Turner, 1996). Quite related to metaphor is the concept of analogy, in which the concept of one mental space is understood in terms of another. What allows for analogical inferences to be effectively made, it has been argued, is the structural similarity between the two spaces (Gentner, Holyoak, & Kokinov, 2001). These approaches to imaginative association based on structural similarity constitute the theoretical grounding for algorithmically implementing the link and shift designs, as well as discourse structures, in the GVR project.

Computational Narrative and Imaginative Discourse Generation

The perspective on computational narrative and imaginative discourse generation here builds upon D. Fox Harrell's approach to producing what he calls "Phantasmal Media." At the broadest level, his practice called imagination computing describes how formal and code-based semantic representations can be leverage for subjective, cultural, critically-aware needs such as in generating narrative or poetry. Based upon cognitive semantics theories above, and using Joseph Goguen's theory of algebraic semiotics to provide a method for formalization (Goguen, 1999), a jointly developed algorithm for conceptual blending called Alloy (Goguen & Harrell, 2009) and a novel framework for computational narrative, Harrell designed and constructed the GRIOT system for implementing generative and interactive narratives (Harrell, 2007). The GRIOT system allows authors to create interactive and generative multimedia works (especially narratives, but including other imaginative discourse forms) where user interaction drives a wide range of guided and structured, but not scripted, eventualities of discourse content and style.

Iconicity and Metaphor in Chinese Characters

In her book (Hiraga, 2005), Masako Hiraga describes a significant relationship between C. S. Peirce's semiotic notion of iconicity and the recent cognitive science results mentioned above. She first reviews Peirce's categorization of signs as icon, index, and symbol, and the division of icon into image, diagram, and metaphor. She then points out that both iconicity and metaphor are based on relationship of similarity, but just along a continuum spanning from perception (concrete visualization) to conception (abstract conceptualization). We found a parallel between Peirce's typology as articulated in Hiraga's framework and the categorization of Chinese character forms. Simple characters like pictographs (e.g., “人” (people), “山” (mountain), and “火” (fire)) correspond to images, and ideographs (e.g., “上” (up), “下” (down)) correspond to diagrams. The former pair refers to similarity in shape while the latter in spatial/graphical structure. Compound characters, which are composites of two or more simple characters, usually generate new meanings by association of paradigms or iconic moments (e.g., “隻” (single) from which one can form “雙” (double), or “災” (disaster), which is composed of “川” (river) and “火” (fire)), many of which are metaphorical icons. This compositionality of meaning has proved inspirational for the visual composition framework in the GVR project.

An Example of GVR

An example work of the GVR project described here presents a type of topographical renku, in which a poetic landscape is interactively co-created by the author-designed system and the user. The landscape is always composed of a series of modular graphic images, which I call topographical tiles, for each of them depicts a topos in our mundane world. The arrangement of the tiles is meaningfully constrained by the structural (i.e., diagrammatic) properties along the edges and then linked perceptually (i.e., imagic) or conceptually (i.e., metaphorical) to one another in response to the user's

successive input. The output is like a topographic map of poetic landscape “picturing” the particular user inclination bouncing back and forth an opposition of organic objects (natural or handcrafted) and modular artifacts (mass-produced or consumerist) that saturate our lives, with animated human figures, or personas, traversing the topography and accumulating their “adaptations” (bodily experiences) to the topoi they have journeyed through.

This section describes the generative algorithms implemented in this work for autonomous and reactive phenomena, the rules applied in producing enduringly transformative personas, and the development of interaction mechanisms to involve motion-sensitive user input.

Semantic Annotation

The initial set of modular graphic images consists of twelve tiles each of which is semantically annotated with perceptual, conceptual, and structural qualities, following the Hiraga’s (and thus Peirce’s) theory mentioned above. These qualities suggest how the tiles can be put together successively in response to user input.

Perceptual and Conceptual Qualities of Tiles

Perceptual qualities include some basic visual attributes, including shape and pattern, together with some labels of other multimedia elements, such as movement and sound. At the conceptual level, tiles are divided into two groups representing the dichotomy of modular and organic objects. The former consists of post-industrial artifacts, which are mostly mass-produced, standardized, and institutionalized. The latter are relatively natural, handcrafted, and local culturally developed (Figure 8.1).

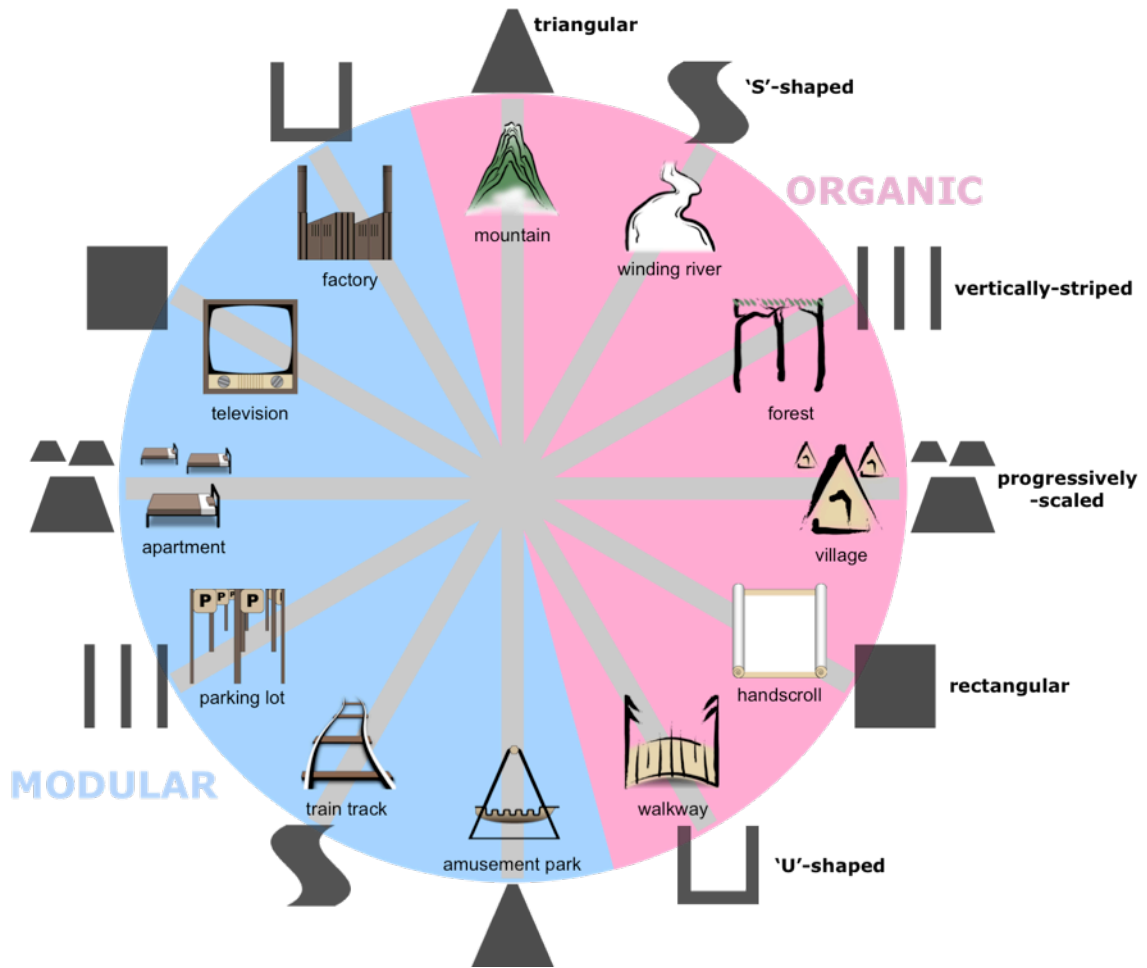


Figure 8.1 A diagram showing the associational (link) and oppositional (shift) relationship between the tiles

Structural Qualities of Tiles

Structural qualities determine how tiles may or may not be conjoined. They implement a combination of cinematic and graphic design conventions considering shot distance and angle, figure/ground relationships, closed/open edges, and the like (Figure 8.2).

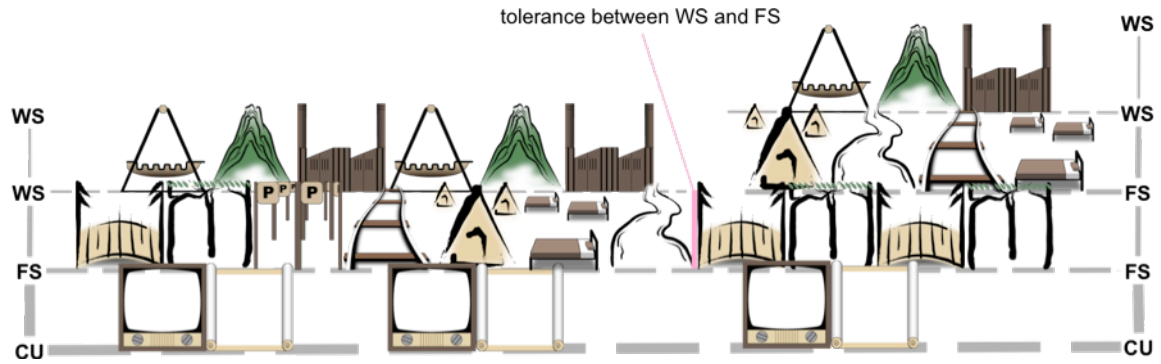


Figure 8.2 A graph showing the edge qualities of the tiles in cinematographic convention: close-up (CU), full shot (FS), wide shot (WS), etc. Please note there exists some tolerance between WS and FS incorporating variability and divergence.

Matching Algorithm

In conceptual metaphor, blending, and analogy theories cited above, elements are mapped from one mental space to another. Building upon insights gained from the Alloy algorithm, Harrell developed an algorithm to find analogous, or “matching,” tiles based upon structural similarity¹⁰ between metadata descriptions of multimedia assets. A general description of the approach follows.

Tiles are put together consecutively. For each pasted tile, the set of possible next tiles is determined using the matching algorithm, which first shortlists possible images based on the weighted structural qualities, and then computes how well the candidates match the previous tile. The degree of a match depends on some combination of the three qualities. The author can specify the relative weightings of the perceptual, structural, and conceptual aspects of images in computing a match. This enables authors to determine their own notions of what comprises a match. In this example work, I set the weightings of the perceptual negative, the conceptual positive, and the structural zero. Hence, a link

¹⁰ What mathematicians call homomorphism.

is a positive-degree match in which the two tiles belong to the same conceptual group (either modular or organic) but differ in shape and pattern. In contrast, a shift is a negative-degree match, also available to the user, and the two tiles would have opposite conceptual group (modular vs. organic) but similar perceptual qualities (e.g., shape). In short, the juxtaposition (either link or shift) of tiles is always constrained by the structural qualities, and the perceptual and conceptual qualities, along with user input, are used to determine whether the juxtaposition is a link or a shift (see Figure 8.3 & 8.4).

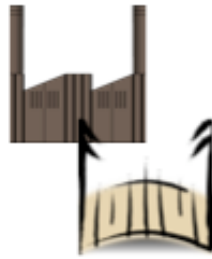


Figure 8.3 A perceptual link (conceptual shift) from factory to walkway, with a structural constraint from WS to FS



Figure 8.4 A conceptual link (perceptual shift) from amusement park to train track, with a structural constraint from WS to FS

The above matching algorithm enables the landscape to extend according to an internal logic. The system always suggests next tiles to be linked either conceptually or perceptually. While the system allows the user to select a direction to extend the landscape, the opposite weightings between perceptual and conceptual qualities force a change in concept to be along with a match in shape or pattern. Therefore, the user would soon realize that a conceptual shift and a perceptual link are coupled. This matching pattern reveals the autonomous transformation of the poetic landscape.

Transformational Rules

Another part of the work is to populate the landscape with context-sensitive human figures in order to charge the output with more evocative metaphors like traditional renku poetry. This transforming figure exemplifies what Harrell has referred to as a polymorphic persona, in this case transforming based upon how it naturalizes in each environment based on appropriate, context specific use or possession of artifacts (Harrell 2010). When the figures traverse the landscape, the personas will change to reflect the human adaptation in different modular or organic topoi. For example, a persona walking in the forest would carry a shotgun. When entering the parking lot, the shotgun-man would be further loaded, becoming a shopping-cart-shotgun-man. In other words, a persona accumulates the track record, and so history, of the walking character, resulting in a transforming persona. This type of transformation is combinatorial, as each of the graphical elements for new artifacts are modular and can be combined with any image of the human. Transformation can occur in other ways, however. A special type of combinatorial relationship is a recursive one. In this type of relationship a graphical element may be nested inside of another one. Finally, we also can transform characters using procedural techniques such as distorting the image using a visual filter, altering the color, or performing a scaling, shearing, or rotating mathematical transformation. Figure 8.5 below depicts a chart of transformational rules for personas inhabiting the poetic landscape.

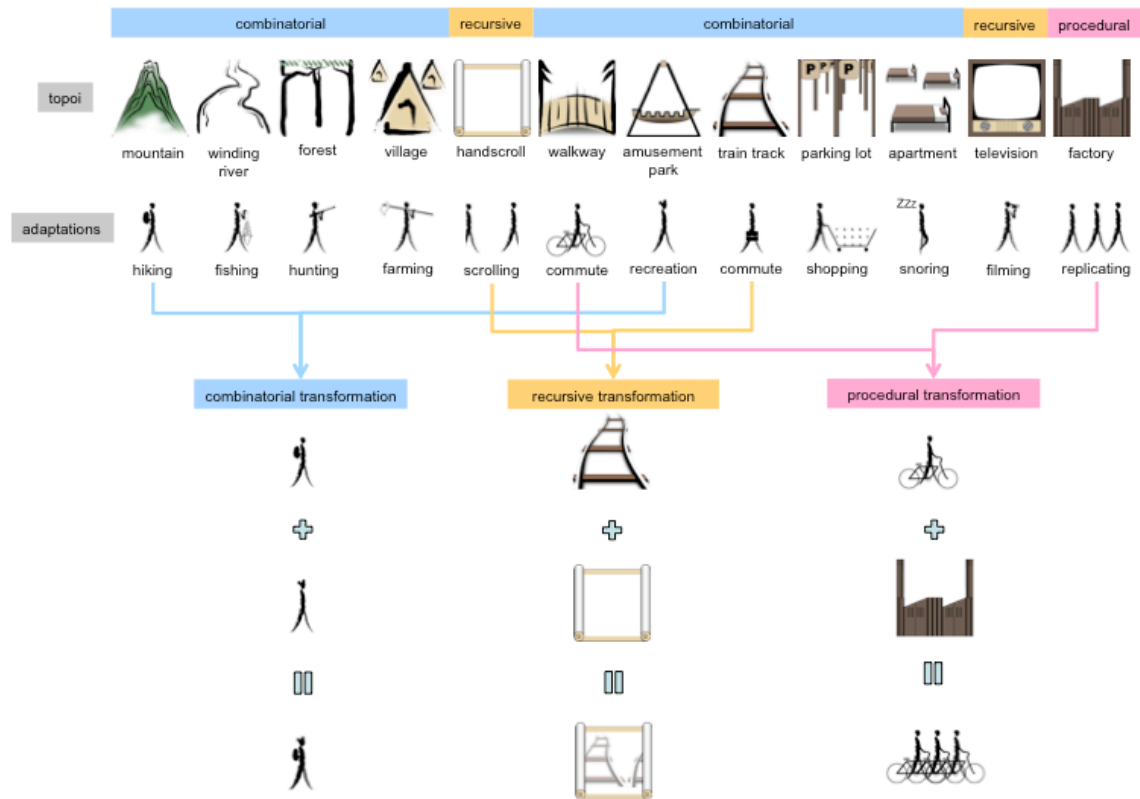


Figure 8.5 Rules for transforming personas and some examples

Transformation according to the rules above brings about liveliness in the walking characters. They seem to metamorphose and evolve autonomously. In case that the system is set to render the link and shift even without user input, for example, by randomly choosing a direction to extend the landscape while idling for a while, the transforming personas would demonstrate the phenomenon of metamorphosis. On the other hand, if the user controls the direction of the traversal, the characters seem to walk in reaction to user input and transform accordingly. In this light, the transforming personas are also reactive.

Interaction Mechanisms

The early implementation of this example work accepted only simple mouse clicks as user selection. After that, we have been developing the interaction mechanism of the work. First, we added the shift-focus effect to the conventional pointing device

input. Second, we are currently exploring the possibilities of implementing other forms of motion-sensitive input.

Shift-Focus Effect

Initially, user input is accepted solely via conventional mouse clicks. In each interaction cycle, a user is provided with possible subsequent tiles in different direction (left, right, up, or down) suggested by the system, and the system would wait for a user click confirming on either one option. In order for an enduring interaction, additional motion-based input is considered. For example, the optional tiles are blurred with respect to how far the mouse pointer moves away from them. When the pointer moves closer to a particular optional tile, the tile would become sharper, imitating the shift-focus effect in cinematography. This simple interactive visual effect is motion sensitive because the degree of blurriness depends on the relative position of the pointer to the tile at any moment. Moreover, this motor-sensory interaction constitutes an elastic anchor for immediate imaginative blends with the user's sensorimotor experience of moving the eyes, or a camera, to wander around and focus on one target at any one time. Further imaginative elaboration of the blend would result in the frame of "decision-making in everyday life." When the pointer rests on a tile, the tile would be clearer than others but still subtly floating around the pointer like something up in the air, invoking a sense of uncertainty because of an experientially based metaphor "unknown is up; known is down" (Lakoff & Johnson, 2003, pp. 20-21). Once the user clicks on a floating tile, it would be pinned down and confirmed, telling that the decision is irreversible.

Finger-Walking Input

Although the shift-focus effect provokes imagination and varies significantly with certain motion parameters, the interaction mechanism is not very enduring because the system would wait indefinitely for a user click so as to confirm a tile. Although the

characters would keep walking without user input, they are animation loops rather than enduring transformation. We started to think of substitutes for mouse clicks. Hence, the touchpad comes into play in our ongoing development, as it is the most accessible motion-sensitive input device for us. Instead of clicking on the touchpad or moving the pointer, we could implement gesture recognition engines to accept various finger actions, such as circling, rubbing, pinching, and others. Meanwhile, this specific motor input should be able to join with the sensory output generated by the system to form a tight motor-sensory connection, making the animated renku an elastic anchor, triggering imaginative elaboration. In the output, the exploration of the landscape and the transformation of the personas are already loaded with rich evocative metaphors. The global one is “life is a journey” back and forth between the post-industrial and the primitive (please see the artists’ statement in (Harrell & Chow, 2009)). In order to echo this main theme, the motion-based input should embody a user’s intention to traverse the landscape. For example, finger walking can be a good candidate, with walking pace affecting the transforming personas’ traversing speed and a turn in walking direction triggering a shift. The act of finger walking ties tightly with the traversal in the landscape, giving rise to new imaginative blends of “a landscape of forking paths” and a life of choices.

Considering New Animation Principles and Variables

When developing this example GVR work, we found that animation phenomena are indispensable in the presentation of generative and interactive visual artworks. Generative rules and interaction mechanisms jointly lead to autonomous transformations and reactive behaviors. The result is a kind of interactive generative animation. Hence, analyzing the example work with respect to the new animation principles and the corresponding variables informs us of room for further development and exploration in the GVR project.

Types of Liveliness

The output of the topographical renku work is an extending poetic landscape populated with walking human characters. On the one hand, both the landscape and characters exhibit emergent liveliness, because the transformations, including the changing brightness of tiles (reflecting the passing time of day) and the walk cycle of characters, are looping. In these loops, no obvious singular source of impulse can be observed. Nor do they bring about irreversible consequence to the topoi. On the other hand, the characters seem to have certain intention as they carry different possessions or change their personas when entering different topoi. These transformative adaptations incite the intentional stance. A user would speculate the reason behind, for example, carrying a shopping cart to the forest. Hence, the work also shows certain extent of intentional liveliness. Apart from that, there is no prominent interactional liveliness, because we do not intend to incorporate obvious causal relationship in the discourse of the work. In summary, the work presents a holistic illusion of life with a bias toward emergent animated patterns.

Levels of Understanding

The human figures and topoi in the topographical renku work are depicted in rudimentary graphic forms, like iconographic or calligraphic styles. The animations, for example, the walk cycle, are reminiscent of our experiences of seeing someone promenading in the park or actually walking by ourselves. When a user selects a direction to extend the landscape, the sensory output blends with the user's sensorimotor experiences of walking out in certain physical environments. In the blend, the connection between the user's finger walking and promenading is compressed. The output is an imaginative navigation back and forth between nostalgic and urban areas. In addition, the aforementioned shift-focus effect adds another immediate blend in which the user "looks at" the possible paths one at a time when standing at a crossroad. The mapping between

pinning down a tile and walking into a chosen path is compressed into a decision made when navigating the poetic virtual space.

Next, the links and shifts in the poetic landscape, together with the transformative adaptations of human figures, provoke further imaginative elaboration on the blends. When the user notices, for example, a character with a shopping cart walking across the boundary between the parking lot and the forest, both of which look similar in shape (see Figure 8.1), the animated output invokes the frame of consumerism and yields a metaphorical narrative in which one inevitably brings along consumer goods of modern societies in his or her everyday life. Since the transformational rules of the personas are thoughtfully designed by the authors, various metaphorical blends emerge from different links/shifts between topoi and the associated changes in character appearance, critically evoking how our social identities transform with subtle traces of memories based upon our environmental and social contexts and utilization of artifacts.

Decelerated Engagement

As described in the section about interaction mechanisms, initial versions of the work largely rely on conventional mouse clicks as input. After suggesting a possible subsequent tile, the system waits indefinitely for a user click. This interaction is halted engagement. To make the engagement continuing, one condition is incorporation of motion-based motor input. Hence, the interactive shift-focus effect is added to consider the relative position of the mouse pointer at any moment and then display corresponding blurriness on the floating tiles. However, the engagement is still decelerated because without a user click no floating tile would be confirmed and the landscape would not extend. Although the human figures keep the walk cycle and some topoi continue animation loops, the whole environment does not evolve. In other words, there is no enduring transformation without user action. To address this inadequacy, the system can be programmed to execute automatic selection of outstanding tiles after a certain while of

idling. Yet, this idea is dismissed because this mechanism violates the essential collaborative ordered structure of linked poetry. Finally, decelerated engagement is regarded as a feature of the topographical renku work.

Degree of Liveness

Stored visual materials in the example GVR work include primarily a set of modular graphic images, comprising all the tiles and human figures. Most transformations, such as the changing brightness of the tiles, the rain or snow effect, the swing of the private ship in the amusement park, and the like are (or will be) implemented algorithmically and dynamically. One exception is the walk cycle of the human personas, which is represented by a sequence of still images (a sprite). Yet, through the transformational rules, including combinatorial, procedural, and recursive methods, the system generates a wide array of possible transformative walking personas. Together with the employment of pseudorandom generation in selection of tiles, deployment of weather effects, and so on, the work is able to cast fairly contingent output, forming a fairly live co-performance between the user and the system.

Summary

The GVR project attempts to extend form-dominant generative approaches in contemporary art to incorporate emotional meanings and evocative contents. The sequentially connected structure and the multitude of emotion-laden topics in renku poetry make it a good candidate to serve as a departure point for our research inquiry. In designing and developing the particular example of GVR works described in this chapter, we construct a semi-structured annotation system of semantic qualities based on Hiraga's framework, and then apply it to an array of visual images. In addition, we adapt the matching algorithm in the GRIOT system to implement new generative system dynamically outputting images constrained on perceptual, structural, and conceptual

bases when accepting user input. Finally, dynamic generation and interaction of images from the system with the user naturally manifest the new animation phenomena I have described, including autonomous transformations, reactive behaviors, metamorphoses, and contingencies.

When designing generative rules and interaction mechanisms for this topographical renku work, I found that the principles of interactive generative animation and the associated variables help not just imbue the artifact with an expanded illusion of life but actually conjoin the form and content. On the one hand, scrutinizing the types of liveliness and the decelerated engagement directs my attention to the form of presentation and interaction. On the other hand, analyzing the levels of imaginative understanding points me to the content. The two sides converge when I examine how the artifact acts as an elastic anchor enabling the motor-sensory connection and triggering imaginative elaboration. In this light, I believe the set of new principles can inform authors or designers a promising way of integrating interactive generative visual artworks with rich evocative and affective meanings. In this new kind of digital media artifacts, content emerges from form, and form hinges on content.

The text and figures of this chapter, in part, contain material that is a reprint of, or has been submitted for publication in (Chow & Harrell, 2010; Harrell & Chow, 2009). The dissertation author was a co-author of this paper.

CHAPTER 9

CONCLUSION

This dissertation introduced a new perspective on animation in today's digitally mediated contexts. This perspective shifts away from existing medium-centric views toward a broader interdisciplinary phenomenon-oriented approach that emphasizes the motor-sensory feedback loops mobilized during human interaction with multimedia artifacts imbued with the new animation phenomena supported by interactive computational generation. This new notion of animation, which I call interactive generative animation, builds upon four areas of knowledge from the vantage points of (1) the observer, (2) the image, (3) the body, and (4) the performer, giving rise to four principles with an array of qualitative variables. I constructed a corpus of media artifacts exemplifying the principles and demonstrating the inclinations of the respective variables, resulting in a novel taxonomy of the new animation paradigm. This taxonomy brings new insights to interaction design from humanistic animation studies and centralizes embodied cognition ideas in designing animation, suggesting designers or developers alternative design perspectives and possibilities. To illustrate the application of the new framework, I described an ongoing generative and interactive visual art project, GVR, and discussed the animation phenomena implemented in the related work. During the project development, the proposed principles and variables reminded us of current inadequacies and suggested possible directions toward achieving all-round liveliness, imaginative understanding, enduring interaction, and divergent computer improvisation. I believe the approach argued for here would benefit the creation of not just digital artworks, but also multimedia artifacts at large, because interactive generative animation makes computational artifacts more “human-familiar” and “life-like.”

Generative and interactive animation expands the meaning of “illusion of life.” The expanded illusion includes images that not only move, but also show reaction, autonomous transformation, metamorphosis, and contingency. These dynamic phenomena of life characterize interactive generative animation, and also reciprocally expand the meaning of “animation.” As I have talked about many meanings of the term at the very beginning of this dissertation, here come several new ones, which resonate with the broad theoretical context of art, poetics, media, and technology.

Animation as an Illusion of Life in the Age of Interactive Computational Generation

In his 1935 seminal essay “The Work of Art in the Age of Mechanical Reproduction,” Walter Benjamin marks the impact of reproduction technologies emerging at that time on the authenticity of the art object (Benjamin, 1968). Mechanical reproduction, he argued, replicates the original and negates its uniqueness in time and space. He eloquently described the phenomenon as “the decay of the aura.” This decay reflects the emergent human desire to bring things “closer,” both “spatially and humanly,” including the one-time sacred art object. This view echoes with my proposal to bring computers “closer” to humans, but the art object is moved from the spiritual end while the computer from the mechanical end. Initially the computer was a tool, instrument, and machine. Interactive multimedia technologies make it more “human” by creating an expanded illusion of life in terms of new animation phenomena.

To Benjamin, film was a new celebrated form of art in that era, because the cinema allows the mass audience to be “unconscious” critics of art by subsuming the cult value of the art object in public exhibition. Perception of film is a middle ground between the concentrated appreciation of painting and the distracted use of architecture. I would add that interactive generative animation is its counterpart in the computer age, and it shifts the original instrumental value of the computer to an imaginative value by immersing human users in enduring and everyday environments, presenting all-round

liveliness in order for their concentrated and diluted attentions, triggering evocative material-based imagination in their minds, and turning them into performers or animators.

Animation as a Simulacrum of Life

While Benjamin asserted that copies problematize the original, Jean Baudrillard further claimed that simulacra displace or dismiss the original (Baudrillard, 1988). By simulacra, he meant the signs in contemporary media context that have no original referents. These media objects seem to imitate something but actually bear no relation to any reality. In fact, a simulacrum is by itself a part of the real, not the imaginary. As Baudrillard's well-known example of Disneyland reveals, the image of the amusement park is not an imaginary world for kids, but instead a simulacrum of the "real" adult world, because the same childishness is everywhere in the societies outside. A cultural critic might say that the ideology of the park is an integral part of the hegemony of American culture. In short, simulacra blur the distinction between the real and the imaginary, and even become part of the real.

In this regard, the new animation paradigm is not just an illusion of life, but also a simulacrum of life. As I have argued, interactive generative animation provokes material-based imagination and destabilizes the opposition between material images (the real) and mental images (the imaginary). People engaging with artifacts of the new animation paradigm see the animated images as "real" phenomena, because the associated motor-sensory experiences are so embodied and evocative. Moreover, this expanded illusion of life is so pervasive in today's digitally mediated contexts that it becomes an integral part of the modern life. The animated multimedia artifacts mentioned in this dissertation apparently imitate people's everyday experiences through all-round liveliness and enduringly changing environments, yet they actually engender unprecedented behaviors and habits, such as web users sending virtual snowflakes with greetings instead of

mailing traditional cards during Christmas, commuters gazing at virtual koi fish anywhere with their hand-held ponds, or people meeting in an art gallery reading motion poems. These new human activities embody the simulation of life supported by interactive generative animation in digital media and bring it back to everyday life. Hence, animation is embodied and realized as a simulacrum.

Animation as a Spiritual-Functional Phenomenon

Animation transcends the instrumental value of the computer, giving way to an imaginative value. Through provoking material-based imagination, animated artifacts give rise to new human behaviors and habits embodying a simulacrum of life. At the heart of this embodiment lies the motor-sensory feedback loop mobilized by generative and interactive systems. Hence, there exists a nuanced interplay between animation and the computer. On the one hand, animation instills the computer with soul and spirit by creating a holistic illusion of life and enduringly changing environments. On the other hand, interactive computational generation enables the sensorimotor functions of animated artifacts. As I argued elsewhere (Chow, 2009), this interplay short-circuits the notions of spirituality and functionality. The two seemingly contrasting terms “spirit” and “function” actually have a subtle and unnoticed link both etymologically and phenomenologically. The Latin origin of “spirit” literally means “breath,” which is a physiological process of taking in air and expelling it, and the ability to breathe is a phenomenal function of living things. Hence, spirituality and functionality jointly describe a phenomenon of life. Obviously the computer and animated artifacts cannot “breathe” literally, yet as this dissertation has demonstrated, they do perform other physiological functions in motor-sensory loops, such as moving in reaction to stimuli. Since interactive generative animation is a simulacrum, the new animation phenomena are “real” and they revitalize the original connection between the spiritual and the functional.

This new spiritual-functional connection “animates” the computer and brings it closer to humans. I call this technological phenomenon spiritual-functional animation. It conceals the complexity and abstraction of the computer, and wraps it with a “skin,” very much like the costumes for marionettes, the shells of automata, the cement of fountains, or the silk screen in shadow puppet theaters. These systems engage the human pursuit to “stage,” but “veil,” technology in order to create spectacles, suspense, and surprise for audiences. In this sense, spiritual-functional animation helps popularize and proliferate computing technology in cultural and economical context. As I argued, the new animation paradigm invites its audiences and users to be animators, puppeteers, or even wizards in computer-mediated co-performance. Through its accessibility, spiritual-functional animation can potentially help adapt every technology-challenged novice into a tinkering hacker. Hence, the phenomenon is not only technological but also cultural, economical, and personal.

Today, animation is omnipresent. The phenomenon of spiritual-functional animation continues absorbing, assimilating, and molding our perception and psychology. It confronts a new uncanny valley, in which we negotiate familiarity and strangeness based on how we situate ourselves relative to those ubiquitous human-familiar digital media artifacts. One day we might see and “believe” animated simulacra in digital media as natural as phenomena in physical reality. Our capacity to hunt or flee in physical space, which is believed to be a major training for human action and perception development in some primordial past (Arnheim, 1974, p. 372), would likely be shifted to techniques to enable us to “crawl” or “surf” in cyberspace.

The text of this chapter, in part, contains material that is a reprint of, or has been submitted for publication in (Chow, 2009). The dissertation author was an author of this paper.

REFERENCES

- Arnheim, R. (1974). *Art and visual perception : a psychology of the creative eye* (New , expanded and rev. ed.). Berkeley: University of California Press.
- Auslander, P. (1999). *Liveness : performance in a mediatized culture*. London ; New York: Routledge.
- Barthes, R. (1973a). *Elements of semiology*. New York: Hill and Wang.
- Barthes, R. (1973b). *Mythologies*. London: Paladin.
- Barthes, R. (1977a). Death of the Author (S. Heath, Trans.) *Image, music, text*. New York: Hill and Wang.
- Barthes, R. (1977b). From Work to Text (S. Heath, Trans.) *Image, music, text* (pp. 220 p., [228] p. of plates). London: Fontana.
- Barthes, R. (1977c). Rhetoric of the Image (S. Heath, Trans.) *Image, music, text*. New York: Hill and Wang.
- Bates, J. (1993). The Nature of Character in Interactive Worlds and The Oz Project. In C. E. Loeffler (Ed.), *Virtual Realities: Anthology of Industry and Culture*.
- Bates, J. (1994). The Role of Emotion in Believable Agents. *Communications of the ACM*, 37(7), 122-125.
- Baudrillard, J. (1988). Simulacra and simulation. In M. Poster (Ed.), *Jean Baudrillard, Selected Writings* (pp. 166-184). Stanford: Stanford University Press.
- Benjamin, W. (1968). The Work of Art in the Age of Mechanical Reproduction (H. Zohn, Trans.) *Illuminations* (pp. 217-251). New York: Schocken Books.
- Bolt, R. A. (1980). "Put-That-There": *Voice and Gesture at the Graphics Interface*. Paper presented at the SIGGRAPH'80.

- Bowker, G. C., & Star, S. L. (1999). *Sorting things out : classification and its consequences*. Cambridge, Mass.: MIT Press.
- Cholodenko, A. (1991a). Introduction. In A. Cholodenko (Ed.), *The Illusion of life : essays on animation* (pp. 9-36). Sydney: Power Publications in association with the Australian Film Commission.
- Cholodenko, A. (2007). Speculations on the Animatic Automaton. In A. Cholodenko (Ed.), *The illusion of life II : more essays on animation* (pp. 486-528). Sydney, N.S.W.: Power Pub.
- Cholodenko, A. (Ed.). (1991b). *The Illusion of life : essays on animation*. Sydney: Power Publications in association with the Australian Film Commission.
- Chow, K. K. N. (2009). The Spiritual-Functional Loop: Animation Redefined in the Digital Age. *Animation: An Interdisciplinary Journal*, 4(1), 77-89.
- Chow, K. K. N., & Harrell, D. F. (2009a). *Active Animation: An Approach to Interactive and Generative Animation for User-Interface Design and Expression*. Paper presented at the Digital Humanities 2009 Conference.
- Chow, K. K. N., & Harrell, D. F. (2009b). *Material-Based Imagination: Embodied Cognition in Animated Images*. Paper presented at the Digital Art and Culture 2009, University of California, Irvine USA.
- Chow, K. K. N., & Harrell, D. F. (2010). *The Generative Visual Renku Project: Integrating Multimedia Semantics, Animation, and User-Interface Design*. Paper presented at the CHI 2010, Atlanta.
- Colebrook, C. (2002a). *Gilles Deleuze*. London ; New York: Routledge.
- Colebrook, C. (2002b). *Understanding Deleuze*. Crows Nest, NSW: Allen & Unwin.
- Crawford, C. (2005). *Chris Crawford on interactive storytelling*. Berkeley, Calif.: New Riders.
- Deleuze, G. (1986). *Cinema*. Minneapolis: University of Minnesota.

- Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass.: MIT Press.
- Dittrich, W., & Lea, S. (1994). Visual Perception of Intentional Motion. *Perception*(23), 253-268.
- Dixon, S. (2007). *Digital performance : a history of new media in theater, dance, performance art, and installation*. Cambridge, Mass.: MIT Press.
- Dourish, P. (2001). *Where the Action Is: The Foundations of Embedded Interaction*. MIT Press.
- Dreyfus, H. L. (1972). *What computers can't do : A critique of artificial reason*. New York: Harper & Row.
- Dreyfus, H. L., Wrathall, M. A., & ebrary Inc. (2006). A companion to phenomenology and existentialism. *Blackwell companions to philosophy* ; 35., from <http://site.ebrary.com/lib/polyu/Doc?id=10158953> Online access from ebrary
- Eco, U. (1976). *A theory of semiotics*. Bloomington: Indiana University Press.
- Eco, U. (1989). *The open work*. Cambridge, Mass.: Harvard University Press.
- Fagone, J. (2008). The Video-Game Programmer Saving Our 21st-Century Souls. *Esquire*.
- Fauconnier, G. (1985). *Mental spaces : aspects of meaning construction in natural language*. Cambridge, Mass.: MIT Press.
- Fauconnier, G. (2001). Conceptual Blending and Analogy. In D. Gentner, K. J. Holyoak & B. N. Kokinov (Eds.), *The analogical mind : perspectives from cognitive science* (pp. 255 - 285). Cambridge, Mass.: MIT Press.
- Fauconnier, G., & Turner, M. (2002). *The way we think : conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Fore, S. (2007). Making Do and Making Meaning: Cultural and Technological Hybridity in Recent Asian Animation. In D. H. Jan Walls, and Emilie Y.Y. Yeh (Ed.), *East-*

West Identities: Globalisation, Localisation, and Hybridisation (pp. 315-328).
Leiden: Brill Academic Publishers.

Furniss, M. (1998). *Art in motion : animation aesthetics*. Sydney, N.S.W.: John Libbey.

Galanter, P. (2006). Generative Art and Rules-based Art. *Vague Terrain*, 3.

Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.). (2001). *The analogical mind : perspectives from cognitive science*. Cambridge, Mass.: MIT Press.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, N.J.: L. Erlbaum.

Glenberg, A. M. (1997). What Memory is For. *Behavioral and Brain Sciences*, 20(1), 1-55.

Goguen, J. (1999). An Introduction to Algebraic Semiotics, with Applications to User Interface Design. In C. L. Nehaniv (Ed.), *Computation for Metaphors, Analogy, and Agents* (pp. 242-291). Berlin: Springer.

Goguen, J., & Harrell, D. F. (2009). Style: A Computational and Conceptual Blending-Based Approach. In S. Argamon & S. Dubnov (Eds.), *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning*. Berlin: Springer-Verlag.

Gombrich, E. H. (2002). *Art and illusion : a study in the psychology of pictorial representation* (6th ed.). London: Phaidon Press.

Goodman, N. (1976). *Languages of art : an approach to a theory of symbols* (2nd ed.). Indianapolis, Ind.: Hakett Publications.

Harrell, D. F. (2007). GRIOT's Tales of Haints and Seraphs: A Computational Narrative Generation System. In P. Harrigan & N. Wardrip-Fruin (Eds.), *Second person : role-playing and story in games and playable media*. Cambridge, Mass.: MIT Press.

- Harrell, D. F. (2009). Toward a Theory of Phantasmal Media: An Imaginative Cognition- and Computation-Based Approach to Digital Media. *CTheory, an international peer-reviewed journal of theory, technology, and culture*.
- Harrell, D. F., & Chow, K. K. N. (2008). *Generative Visual Renku: Linked Poetry Generation with the GRIOT System*. Paper presented at the Visionary Landscapes: Electronic Literature Organization 2008 Conference.
- Harrell, D. F., & Chow, K. K. N. (2009). Generative Visual Renku: Poetic Multimedia Semantics with the GRIOT System. *Hyperrhiz: New Media Cultures*(Special Issue: Visionary Landscapes).
- Harrell, D. F., & Zhu, J. (2009). *Agency Play: Dimensions of Agency for Interactive Narrative Design*. Paper presented at the Proceeding of the AAAI 2009 Spring Symposium on Narrative Intelligence II.
- Higginson, W. J. (1996). *The haiku seasons : poetry of the natural world* (1st ed.). Tokyo: Kodansha International.
- Hiraga, M. (2005). *Metaphor and iconicity : a cognitive approach to analysing texts*. Houndmills, Basingstoke, Hampshire ; New York: Palgrave Macmillan.
- Hutchins, E. (2005). Material Anchors for Conceptual Blends. *Journal of Pragmatics*, 37(10).
- Johnson, S. (1997). *Interface culture : how new technology transforms the way we create and communicate* (1st ed.). [San Francisco]: HarperEdge.
- Kaplin, S. (2001). A Puppet Tree: A Model for the Field of Puppet Theatre. In J. Bell (Ed.), *Puppets, masks, and performing objects* (pp. 18-25). Cambridge, Mass. ; London: MIT Press.
- Klein, N. M. (1993). *Seven minutes : the life and death of the American animated cartoon*. London ; New York: Verso.
- Klein, N. M. (2004). *The Vatican to Vegas : a history of special effects*. New York: New Press.

- Kleist, H. v. (1810). On the Marionette Theatre. Retrieved from <http://www.southerncrossreview.org/9/kleist.htm>
- Kurtenbach, G., & Hulteen, E. A. (1990). Gestures in Human-Computer Communication. In B. Laurel & S. J. Mountford (Eds.), *The Art of human-computer interface design* (pp. 309-317). Reading, Mass.: Addison-Wesley Pub. Co.
- Lakoff, G. (1987). *Women, fire, and dangerous things : what categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh : the embodied mind and its challenge to Western thought*. New York: Basic Books.
- Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., & Turner, M. (1989). *More than cool reason : a field guide to poetic metaphor*. Chicago: University of Chicago Press.
- Lamarre, T. (2009). *The anime machine : a media theory of animation*. Minneapolis: University of Minnesota Press.
- Lanier, J. (1995). Agents of Alienation. *Journal of Consciousness Studies*, 2, 76-81.
- Lewis, G. E. (2000). Too Many Notes: Computers, Complexity and Culture in Voyager. *Leonardo Music Journal*, 10, 33-39.
- Lunenfeld, P. (Ed.). (1999). *The digital dialectic : new essays on new media*. Cambridge, Mass.: MIT Press.
- Manovich, L. (2001). *The Language of New Media*. Cambridge, MA: MIT Press.
- Mateas, M. (2004). A Preliminary Poetics for Interactive Drama and Games. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First person : new media as story, performance, and game*. Cambridge, Mass.: MIT Press.
- Mathews, H., & Brotchie, A. (Eds.). (1998). *Oulipo compendium*. London: Atlas Press.

- Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge & Kegan Paul.
- Mitchell, W. J. T. (1986). *Iconology : image, text, ideology*. Chicago: University of Chicago Press.
- Moran, D. (2000). *Introduction to phenomenology*. London: Routledge.
- Moritz, W. (1997). Digital Harmony: The Life of John Whitney, Computer Animation Pioneer. *Animation World Magazine*, 2(5). Retrieved from <http://www.awn.com/mag/issue2.5/2.5pages/2.5moritzwhitney.html>
- Murray, J. (1997). *Hamlet on the Holodeck*. Cambridge, MA: MIT Press.
- Napier, S. J. (2005). *Anime from Akira to Howl's moving castle : experiencing contemporary Japanese animation* (Updated ed.). New York, N.Y.: Palgrave Macmillan.
- Nielsen, J. (2000). *Designing web usability*. Indianapolis, Ind.: New Riders.
- Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books Inc.
- Norman, D. A. (2002). *The Design of Everyday Things*. New York: Basic Books Inc.
- Ono, K. (2007). The Long Flight of Manga and Anime: The History of Comics and Animation in Japan. In A. Cholodenko (Ed.), *The illusion of life II : more essays on animation* (pp. 99-118). Sydney, N.S.W.: Power Pub.
- Penny, S. (1996). From A to D and back again: The emerging aesthetics of Interactive Art. *Leonardo Electronic Almanac*(April 1996).
- Reiniger, L. (1970). *Shadow theatres and shadow films*. London, New York,: Batsford; Watson-Guption Publications.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain : how our minds share actions and emotions*. Oxford ; New York: Oxford University Press.

- Rohrer, J. (2007, 2007). What I was trying to do with Passage, 2010, from <http://hcsoftware.sourceforge.net/passage/statement.html>
- Russett, R., & Starr, C. (1988). *Experimental animation : origins of a new art* (Rev. ed.). New York ; London: Da Capo.
- Saussure, F. d. (1983). *Course in general linguistics* (R. Harris, Trans.). London: Duckworth.
- Schaffer, W. (2007). Animation 1: The Control-Image. In A. Cholodenko (Ed.), *The illusion of life II : more essays on animation* (pp. 456-485). Sydney, N.S.W.: Power Pub.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual Causality and Animacy. *Trends in Cognitive Sciences*, 4(8).
- Shneiderman, B. (2003 [1983]). Direct Manipulation: A Step beyond Programming Languages. In N. Wardrip-Fruin & N. Montfort (Eds.), *The new media reader* (pp. 486-498). Cambridge, Mass. ; London: MIT Press.
- Shneiderman, B., & Maes, P. (1997). Direct manipulation versus interface agents. *Interactions*, 4(6), 42-61.
- Sobchack, V. C. (2000). *Meta-morphing : visual transformation and the culture of quick-change*. Minneapolis: University of Minnesota Press.
- Suchman, L. A. (2007). *Human-machine reconfigurations : plans and situated actions* (2nd ed.). Cambridge ; New York: Cambridge University Press.
- Terpak, F. (2001). AUTOMATA (BATTERIES NOT INCLUDED). In B. M. Stafford, F. Terpak, I. Poggi & J. Paul Getty Museum. (Eds.), *Devices of wonder : from the world in a box to images on a screen* (pp. 266-274). Los Angeles, Calif.: Getty Research Institute.
- Thomas, F., & Johnston, O. (1984). *Disney animation : the illusion of life* (Popular ed.). New York: Abbeville Press.

- Tillis, S. (2001). The Art of Puppetry in the Age of Media Production. In J. Bell (Ed.), *Puppets, masks, and performing objects* (pp. 172-185). Cambridge, Mass. ; London: MIT Press.
- Turkle, S. (1984). *The second self : computers and the human spirit*. New York: Simon and Schuster.
- Turner, M. (1996). *The literary mind*. New York: Oxford University Press.
- Ward, P. (2000). Defining "Animation": The Animated Film and the Emergence of the Film Bill. *Scope: An Online Journal of Film and TV Studies*.
- Ward, P. (2003). Animation Studies, Disciplinarity and Discursivity. *Reconstruction: Studies in Contemporary Culture*, 3(2).
- Wells, P. (1998). *Understanding animation*. London ; New York: Routledge.
- Wells, P. (2002). *Animation and America*. New Brunswick, N.J.: Rutgers University Press.
- Wexelblat, A. (1994). *A feature-based approach to continuous-gesture analysis*. Massachusetts Institute of Technology.
- Wiener, N. (1961). *Cybernetics : or, control and communication in the animal and the machine* (2d ed.). New York: M.I.T. Press.
- Wu, H. (1996). *The double screen : medium and representation in Chinese painting*. London: Reaktion Books.
- Wyche, S. P., Caine, K. E., Davison, B., Arteaga, M., & Grinter, R. E. (2008). *Sun Dial: Exploring Techno-Spiritual Design through a Mobile Islamic Call to Prayer Application*. Paper presented at the CHI 2008, Florence, Italy.
- Zhu, J., & Harrell, D. F. (2009). *The Artificial Intelligence (AI) Hermeneutic Network: Toward an Approach to Analysis and Design of Intentional Systems*. Paper presented at the Digital Humanities 2009.